



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **09134364 A**(43) Date of publication of application: **20.05.97**

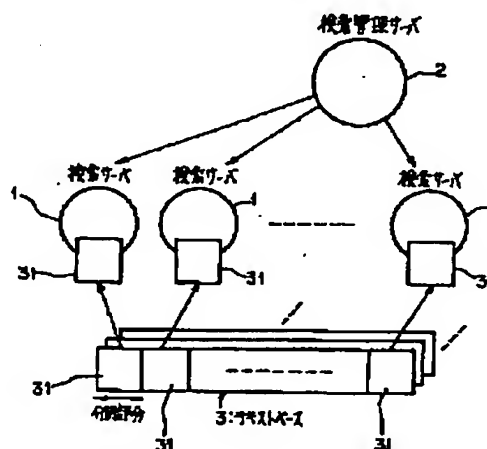
(51) Int. Cl.

G06F 17/30(21) Application number: **07290218**(71) Applicant: **FUJITSU LTD**(22) Date of filing: **08.11.95**(72) Inventor: **OE KAZUICHI****(54) INFORMATION RETRIEVAL SYSTEM****(57) Abstract:**

PROBLEM TO BE SOLVED: To quickly retrieve information without requiring special hardware and without being affected by the size of a text base to be retrieved even in the case of simultaneously retrieving plural text bases.

SOLUTION: The information retrieving system is provided with plural retrieving servers 1 for executing retrieving processing and a retrieval managing server 2 for managing the operation of these servers 1. The server 2 divides text bases 3 to be retrieved and the relational information of respective text bases 3 and allocates the divided parts 31 of the text base 3 and relational information corresponding to respective divided parts 31 to a part or all of the plural retrieving servers 1 in a pair. Information retrieval to the divided parts 31 of the text bases 3 allocated to respective servers 1 by the server 2 is independently executed in parallel.

COPYRIGHT: (C)1997,JPO



Best Available Copy

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-134364

(43) 公開日 平成9年(1997)5月20日

(51) Int.Cl.⁴

G 0 6 F 17/30

識別記号

庁内整理番号

P I

G 0 6 F 15/40

技術表示箇所

3 1 0 B

審査請求 未請求 請求項の数17 O L (全 34 頁)

(21) 出願番号 特願平7-290218

(22) 出願日 平成7年(1995)11月8日

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 大江 和一

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(74) 代理人 弁理士 真田 有

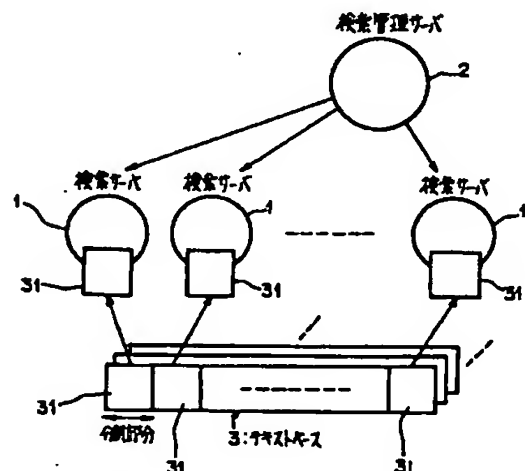
(54) 【発明の名称】 情報検索システム

(57) 【要約】

【課題】 情報検索システムに関し、複数のテキストベースを同時に検索する場合でも、特別なハードウェアを必要とせず、且つ、検索対象テキストベースの大きさにも影響されずに高速に情報検索を行なう。

【解決手段】 検索処理を行なう複数の検索サーバ1と、これら複数の検索サーバ1の動作を管理するための検索管理サーバ2とをそなえ、検索管理サーバ2を、検索対象テキストベース3とこのテキストベース3についての関連情報とをそれぞれ分割し、このテキストベース3の分割部分31とその分割部分31に対応する関連情報とを組にして複数の検索サーバ1のうちの一部または全部に割り当てるように構成するとともに、複数の検索サーバ1を、それぞれ、検索管理サーバ2によって割り当てられたテキストベース3の分割部分31に対する情報検索を並列かつ独立に行なうように構成する。

本発明の原理ブロック図



【特許請求の範囲】

【請求項1】 検索処理を行なう複数の検索サーバと、
該複数の検索サーバの動作を管理するための検索管理サ
ーバとをそなえ、

該検索管理サーバが、検索対象テキストベースと該テキ
ストベースについての関連情報とをそれぞれ分割し、該
テキストベースの分割部分とその分割部分に対応する関
連情報とを組にして前記複数の検索サーバのうちの一部
または全部に割り当てるように構成されるとともに、
該複数の検索サーバが、それぞれ、該検索管理サーバに
よって割り当てられた該テキストベースの分割部分に対
する情報検索を並列かつ独立に行なうように構成されて
いることを特徴とする、情報検索システム。

【請求項2】 該検索管理サーバが、
検索サーバ毎に、検索を担当しているテキストベース名
とそのテキストベースについて検索を担当している分割
部分のデータ量とを含む検索サーバ情報を管理する検索
サーバ管理機能と、

検索対象テキストベース毎に、そのテキストベース名、
そのテキストベース実体を格納するテキストファイル
名、そのテキストベースについての関連情報を格納する
関連ファイル名を管理するテキストベース管理機能と、
テキストベースの検索を新たに始める際に、当該テキ
ストベースのサイズ情報と該検索サーバ管理機能で管理
されている検索サーバ情報から得られる各検索サーバの負
荷状況とに基づいて当該テキストベースの分割数を決定
する分割数決定機能と、

該分割数決定機能により決定された分割数と前記の各検
索サーバの負荷状況とに応じて、当該テキストベースと
その関連情報とを、負荷の軽い検索サーバから順に割り
当てる検索サーバ割当機能と、

該検索サーバ割当機能による割当結果に基づいて、当該
テキストベースの分割部分とその関連情報とを割当先の
検索サーバへロードするためのローディング指示を、該
テキストベース管理機能で管理されているテキストファ
イル名および関連情報ファイル名とともに割当先の各検
索サーバへ送るローディング指示機能とを有しているこ
とを特徴とする、請求項1記載の情報検索システム。

【請求項3】 各検索サーバが、

該検索管理サーバからのローディング指示を受けると、
該ローディング指示に伴うテキストファイル名および関
連情報ファイル名によってそれぞれ指定されるテキス
トベースの分割部分およびその関連情報を読み込むファ
イルアクセス機能と、

検索要求を受けると、該ファイルアクセス機能により読
み込んだテキストベースの分割部分に対する情報検索を
行なう検索機能と、

該検索機能による検索結果を検索要求元に返す検索結
果送出機能とを有していることを特徴とする、請求項2記
載の情報検索システム。

【請求項4】 該検索管理サーバが、

各検索サーバ毎に、検索を担当しているテキストベース
名、そのテキストベースについて検索を担当している分
割部分のデータ量、使用可能メモリ量を含む検索サーバ
情報を管理する検索サーバ管理機能と、

検索対象テキストベース毎に、そのテキストベース名、
そのテキストベース実体を格納するテキストファイル
名、そのテキストベースについての関連情報を格納する
関連ファイル名、使用すべき検索アルゴリズム情報を管
理するテキストベース管理機能と、

テキストベースの検索を新たに始める際に、当該テキ
ストベースのサイズ情報と該テキストベース管理機能で管
理されている検索アルゴリズム情報と該検索サーバ管理
機能で管理されている各検索サーバの使用可能メモリ量
とに基づいて当該テキストベースの分割数を決定する分
割数決定機能と、

該分割数決定機能により決定された分割数と前記の各検
索サーバの使用可能メモリ量とに応じて、当該テキス
トベースとその関連情報とを、使用可能メモリ量の大きい
検索サーバから順に割り当てる検索サーバ割当機能と、
該検索サーバ割当機能による割当結果に基づいて、当該
テキストベースの分割部分とその関連情報とを割当先の
検索サーバへロードするためのローディング指示を、該
テキストベース管理機能で管理されているテキストファ
イル名、関連情報ファイル名および検索アルゴリズム情
報とともに割当先の各検索サーバへ送るローディング指
示機能とを有していることを特徴とする、請求項1記載
の情報検索システム。

【請求項5】 各検索サーバが、

該検索管理サーバからのローディング指示を受けると、
該ローディング指示に伴うテキストファイル名および関
連情報ファイル名によってそれぞれ指定されるテキス
トベースの分割部分およびその関連情報を読み込むファ
イルアクセス機能と、

複数種類の検索アルゴリズムをそなえ、検索要求を受け
ると、該検索管理サーバからの検索アルゴリズム情報に
より指定される検索アルゴリズムを用いて、該ファイル
アクセス機能により読み込んだテキストベースの分割部
分に対する情報検索を行なう検索機能と、

該検索機能による検索結果を検索要求元に返す検索結
果送出機能とを有していることを特徴とする、請求項4記
載の情報検索システム。

【請求項6】 ユーザプロセスからの検索要求を受け付
けるユーザ管理サーバをそなえ、

該ユーザ管理サーバが、ユーザプロセスからの検索要求
に応じて、検索対象テキストベースに対する検索を行な
っている検索サーバ群情報の照会を該検索管理サーバに
行ない、その照会結果に基づいて実際の検索要求を各検
索サーバに対して直接的に行なうように構成されている
ことを特徴とする、請求項1記載の情報検索システム。

【請求項7】 該ユーザ管理サーバが、ユーザプロセスからの検索要求を受け付ける検索要求受付機能と、

該検索要求受付機能により受け付けた検索要求に応じて、当該検索要求の検索対象テキストベースに対する検索処理を割り当てられている検索サーバ群情報を該検索管理サーバに照会する検索サーバ照会機能と、該検索サーバ照会機能により該検索管理サーバから得られた検索サーバ群情報に基づいて、当該検索要求の検索対象テキストベースに対する検索処理を割り当てられている各検索サーバに対して直接的に検索要求を送出する検索要求機能と、

該検索要求機能の検索要求に対する各検索サーバからの応答をマージし検索結果として検索要求元のユーザプロセスに転送する検索結果転送機能とを有していることを特徴とする、請求項6記載の情報検索システム。

【請求項8】 該検索管理サーバが、検索サーバ毎に、検索を担当しているテキストベース名とそのテキストベースについて検索を担当している分割部分のデータ量とを含む検索サーバ情報を管理する検索サーバ管理機能と、

検索対象テキストベースのローディング先である検索サーバ群に関するローディング情報を管理するローディング情報管理機能と、

検索対象テキストベース毎に、そのテキストベース名、そのテキストベース実体を格納するテキストファイル名、そのテキストベースについての関連情報を格納する関連ファイル名を管理するテキストベース管理機能と、該ユーザ管理サーバから受けた照会要求の検索対象テキストベースが未だロードされていない場合、当該テキストベースのサイズ情報と該検索サーバ管理機能で管理されている検索サーバ情報から得られる各検索サーバの負荷状況とに基づいて当該テキストベースの分割数を決定する分割数決定機能と、

該分割数決定機能により決定された分割数と前記の各検索サーバの負荷状況とに応じて、当該テキストベースとその関連情報とを、負荷の軽い検索サーバから順に割り当てる検索サーバ割当機能と、

該検索サーバ割当機能による割当結果に基づいて、当該テキストベースの分割部分とその関連情報とを割当先の検索サーバへロードするためのローディング指示を、該テキストベース管理機能により得られるテキストファイル名および関連情報ファイル名とともに割当先の各検索サーバへ送るローディング指示機能と、

該検索サーバ割当機能により当該テキストベースを割り当てられた検索サーバ群に関するローディング情報を、検索サーバ群情報として該ユーザ管理サーバへ転送するローディング情報転送機能とを有していることを特徴とする、請求項7記載の情報検索システム。

【請求項9】 該分割数決定機能により決定された分割

数分の検索サーバを該検索サーバ割当機能により確保できない場合には、該分割数決定機能により決定された当該テキストベースの分割数を見直し、その分割数に応じて該検索サーバ割当機能、該ローディング指示機能および該ローディング情報転送機能による処理を行なうことを特徴とする、請求項8記載の情報検索システム。

【請求項10】 該ユーザ管理サーバから受けた照会要求の検索対象テキストベースが既にロードされている場合、該ローディング情報転送機能が、該ローディング情報管理機能で管理されている当該検索対象テキストベースについてのローディング情報を、検索サーバ群情報として該ユーザ管理サーバへ転送することを特徴とする、請求項8または請求項9に記載の情報検索システム。

【請求項11】 該ユーザ管理サーバから受けた照会要求の検索対象テキストベースが既にロードされている場合、該分割数決定機能が、当該テキストベースのサイズ情報と該検索サーバ管理機能で管理されている検索サーバ情報から得られる各検索サーバの負荷状況とに基づいて当該テキストベースの分割数を決定し、

該複数の検索サーバの中から所定値よりも負荷の軽い検索サーバを該分割数決定機能により決定された分割数だけ確保できる場合、該検索サーバ割当機能が、前記所定値よりも負荷の軽い検索サーバを当該分割数だけ選択して当該テキストベースとその関連情報とを割り当て、該ローディング指示機能が、該検索サーバ割当機能による割当結果に基づいて、当該テキストベースの分割部分とその関連情報とを割当先の検索サーバへロードするためのローディング指示を、該テキストベース管理機能により得られるテキストファイル名および関連情報ファイル名とともに割当先の各検索サーバへ送り、該ローディング情報転送機能が、該検索サーバ割当機能により当該テキストベースを割り当てられた検索サーバ群に関するローディング情報を、検索サーバ群情報として該ユーザ管理サーバへ転送する一方、

該複数の検索サーバの中から前記所定値よりも負荷の軽い検索サーバを該分割数決定機能により決定された分割数だけ確保できない場合、該ローディング情報転送機能が、該ローディング情報管理機能で管理されている当該検索対象テキストベースについてのローディング情報を、検索サーバ群情報として該ユーザ管理サーバへ転送することを特徴とする、請求項8または請求項9に記載の情報検索システム。

【請求項12】 該検索サーバが、該検索管理サーバからのローディング指示を受けると、該ローディング指示に伴うテキストファイル名および関連情報ファイル名によってそれぞれ指定されるテキストベースの分割部分およびその関連情報を読み込むファイルアクセス機能と、

該ユーザ管理サーバから検索要求を受けると、該ファイルアクセス機能により読み込んだテキストベースの分割

部分に対する情報検索を行なう検索機能と、
該検索機能による検索結果を該ユーザ管理サーバに返す
検索結果送出機能とを有していることを特徴とする、請
求項 8～請求項 11 のいずれかに記載の情報検索システ
ム。

【請求項 13】 該検索管理サーバが、
検索サーバ毎に、検索を担当しているテキストベース
名、そのテキストベースについて検索を担当している分
割部分のデータ量、使用可能メモリ量を含む検索サーバ
情報を管理する検索サーバ管理機能と、
検索対象テキストベースのローディング先である検索サ
ーバ群に関するローディング情報を管理するローディン
グ情報管理機能と、
検索対象テキストベース毎に、そのテキストベース名、
そのテキストベース実体を格納するテキストファイル
名、そのテキストベースについての関連情報を格納する
関連ファイル名、使用すべき検索アルゴリズム情報を管
理するテキストベース管理機能と、
該ユーザ管理サーバから受けた照会要求の検索対象テキ
ストベースが未だロードされていない場合、当該テキス
トベースのサイズ情報と該テキストベース管理機能で管
理されている検索アルゴリズム情報と該検索サーバ管理
機能で管理されている各検索サーバの使用可能メモリ量
とに基づいて当該テキストベースの分割数を決定する分
割数決定機能と、
該分割数決定機能により決定された分割数と前記の各検
索サーバの使用可能メモリ量とに応じて、当該テキス
トベースとその関連情報とを、使用可能メモリ量の大きい
検索サーバから順に割り当てる検索サーバ割当機能と、
該検索サーバ割当機能による割当結果に基づいて、当該
テキストベースの分割部分とその関連情報とを割当先の
検索サーバへロードするためのローディング指示を、該
テキストベース管理機能により得られるテキストファイ
ル名および関連情報ファイル名とともに割当先の各検索
サーバへ送るローディング指示機能と、
該検索サーバ割当機能により当該テキストベースを割り
当てられた検索サーバ群に関するローディング情報を、
検索サーバ群情報として該ユーザ管理サーバへ転送する
ローディング情報転送機能とを有していることを特徴と
する、請求項 7 記載の情報検索システム。

【請求項 14】 該分割数決定機能により決定された分
割数分の検索サーバを該検索サーバ割当機能により確保
できない場合には、該分割数決定機能により決定された
当該テキストベースの分割数を見直し、その分割数に応
じて該検索サーバ割当機能、該ローディング指示機能お
よび該ローディング情報転送機能による処理を行なうこ
とをことを特徴とする、請求項 13 記載の情報検索シス
テム。

【請求項 15】 該ユーザ管理サーバから受けた照会要
求の検索対象テキストベースが既にロードされている場

合、該ローディング情報転送機能が、該ローディング情
報管理機能で管理されている当該検索対象テキストベー
スについてのローディング情報を、検索サーバ群情報と
して該ユーザ管理サーバへ転送することを特徴とする、
請求項 13 または請求項 14 に記載の情報検索システ
ム。

【請求項 16】 該ユーザ管理サーバから受けた照会要
求の検索対象テキストベースが既にロードされている場
合、該分割数決定機能が、当該テキストベースのサイズ
情報と該テキストベース管理機能で管理されている検索
アルゴリズム情報と該検索サーバ管理機能で管理されて
いる各検索サーバの使用可能メモリ量とに基づいて当該
テキストベースの分割数を決定し、
該複数の検索サーバの中から所定値よりも使用可能メモ
リ量の大きい検索サーバを該分割数決定機能により決定
された分割数だけ確保できる場合、該検索サーバ割当機
能が、前記所定値よりも使用可能メモリ量の大きい検索
サーバを当該分割数だけ選択して当該テキストベースと
その関連情報とを割り当て、該ローディング指示機能
が、該検索サーバ割当機能による割当結果に基づいて、
当該テキストベースの分割部分とその関連情報とを割当
先の検索サーバへロードするためのローディング指示
を、該テキストベース管理機能により得られるテキスト
ファイル名、関連情報ファイル名および検索アルゴリズム
情報とともに割当先の各検索サーバへ送り、該ローデ
ィング情報転送機能が、該検索サーバ割当機能により当
該テキストベースを割り当てられた検索サーバ群に関する
ローディング情報を、検索サーバ群情報として該ユー
ザ管理サーバへ転送する一方、
該複数の検索サーバの中から前記所定値よりも使用可能
メモリ量の大きい検索サーバを該分割数決定機能により
決定された分割数だけ確保できない場合、該ローディン
グ情報転送機能が、該ローディング情報管理機能で管理
されている当該検索対象テキストベースについてのロー
ディング情報を、検索サーバ群情報として該ユーザ管理
サーバへ転送することを特徴とする、請求項 13 または
請求項 14 に記載の情報検索システム。

【請求項 17】 該検索サーバが、
該検索管理サーバからのローディング指示を受けると、
該ローディング指示に伴うテキストファイル名および関
連情報ファイル名によってそれぞれ指定されるテキスト
ベースの分割部分およびその関連情報を読み込むファイ
ルアクセス機能と、

複数種類の検索アルゴリズムをそなえ、該ユーザ管理サ
ーバから検索要求を受けると、該検索管理サーバからの
検索アルゴリズム情報により指定される検索アルゴリ
ズムを用いて、該ファイルアクセス機能により読み込んだ
テキストベースの分割部分に対する情報検索を行なう検
索機能と、

該検索機能による検索結果を該ユーザ管理サーバに返す

検索結果送出機能とを有していることを特徴とする、請求項13～請求項16のいずれかに記載の情報検索システム。

【発明の詳細な説明】

【0001】（目次）

発明の属する技術分野

従来の技術

発明が解決しようとする課題

課題を解決するための手段（図1）

発明の実施の形態

（a）情報検索システム全体の構成説明（図2～図10）

（b）ユーザ管理サーバの動作説明（図11～図14）

（c）検索管理サーバの動作説明（図15～図19）

（d）検索サーバの動作説明（図20）

発明の効果

【0002】

【発明の属する技術分野】本発明は、情報検索システムに関し、特に、ファイルシステムなどの各種データベースにおいて所望の情報を高速に検索する、情報検索システムに関する。

【0003】

【従来の技術】周知のように、情報検索システムにおける情報検索処理は、計算機の資源〔CPU（中央処理装置）、メモリ、I/O（入出力装置）など〕を大量に使用（専有）して行なわれるため、計算機の負荷が非常に重くなり、検索速度が低下してしまう場合がある。

【0004】そこで、従来の情報検索システムでは、情報検索専用のハードウェアなどを付加したりして計算機の負荷を軽減することによって、検索速度の向上を図っている。

【0005】

【発明が解決しようとする課題】しかしながら、上述のように情報検索専用のハードウェアを付加した場合でも、検索対象となるデータベース（例えば、新聞や特許文書などのテキスト文書を保持するテキストベース）が大規模になると、その検索能力の限界から、やはり、検索速度は低下してしまう。また、このように情報検索専用のハードウェアを付加することは、コストの面でも非常に不利である。

【0006】さらに、一般に従来の情報検索システムは、基本的に逐次システム、つまり、テキストベース内の情報と検索すべき情報との比較・判定処理を逐次に行なうことによって検索を行なうシステムであるため、例えば、複数のユーザが同時にシステムを利用して、複数のテキストベースを検索しなければならない場合には、検索速度が大幅に低下してしまうという課題もある。

【0007】本発明は、このような課題に鑑み創案されたもので、検索対象テキストベースを分割し、各分割部

分に対する情報検索をそれぞれ複数の検索サーバが並列かつ独立に行なうことによって、複数のテキストベースを同時に検索する場合でも、特別なハードウェアを必要とせず、且つ、そのテキストベースの大きさにも影響されずに高速に情報検索を行なえるようにした、情報検索システムを提供することを目的とする。

【0008】

【課題を解決するための手段】図1は本発明の原理ブロック図で、この図1において、1は検索処理を行なう複数の検索サーバ、2は各検索サーバ1の動作を管理するための検索管理サーバ、3は各検索サーバ1による検索処理対象となる検索対象テキストベースである。

【0009】そして、この場合、検索管理サーバ2は、検索対象テキストベース3とこのテキストベース3についての関連情報とをそれぞれ分割し、テキストベース3の分割部分31とその分割部分31に対応する関連情報とを組にして各検索サーバ1のうちの全部または全部に割り当てるように構成され、各検索サーバ1は、それぞれ、検索管理サーバ2によって割り当てられたテキストベース3の分割部分31に対する情報検索を並列かつ独立に行なうように構成される（請求項1）。

【0010】具体的に、上述の検索管理サーバ2は、以下のような各機能を有して構成される（請求項2）。

（1a）各検索サーバ1毎に、検索を担当しているテキストベース名とそのテキストベース3について検索を担当している分割部分31のデータ量とを含む検索サーバ情報を管理する検索サーバ管理機能。

【0011】（2a）検索対象テキストベース3毎に、そのテキストベース名、そのテキストベース実体を格納するテキストファイル名、そのテキストベース3についての関連情報を格納する関連ファイル名を管理するテキストベース管理機能。

（3a）テキストベース3の検索を新たに始める際に、テキストベース3のサイズ情報と上記（1a）の検索サーバ管理機能で管理されている検索サーバ情報から得られる各検索サーバ1の負荷状況とに基づいてテキストベース3の分割数を決定する分割数決定機能。

【0012】（4a）上記（3a）の分割数決定機能により決定された分割数と各検索サーバ1の負荷状況とに応じて、テキストベース3とその関連情報とを、各検索サーバ1のうち負荷の軽い検索サーバ1から順に割り当てる検索サーバ割当機能。

（5a）上記（4a）の検索サーバ割当機能による割当結果に基づいて、テキストベース3の分割部分31とその関連情報とを割当先の検索サーバ1へロードするためのローディング指示を、上記（2a）のテキストベース管理機能で管理されているテキストファイル名および関連情報ファイル名とともに割当先の検索サーバ1へ送るローディング指示機能。

【0013】一方、各検索サーバ1は、以下のような各

10

20

30

40

50

機能を有して構成される（請求項3）。

（1b）検索管理サーバ2からのローディング指示を受けると、このローディング指示に伴うテキストファイル名および関連情報ファイル名によってそれぞれ指定されるテキストベース3の分割部分31およびその関連情報を読み込むファイルアクセス機能。

【0014】（2b）検索要求を受けると、上記（1b）のファイルアクセス機能により読み込んだテキストベース3の分割部分31に対する情報検索を行なう検索機能。

（3b）上記（2b）の検索機能による検索結果を検索要求元に返す検索結果送出機能。なお、上述の検索管理サーバ2は、以下の各機能を有して構成してもよい（請求項4）。

【0015】（1c）各検索サーバ1毎に、検索を担当しているテキストベース名、そのテキストベース3について検索を担当している分割部分31のデータ量、使用可能メモリ量を含む検索サーバ情報を管理する検索サーバ管理機能。

（2c）検索対象テキストベース3毎に、そのテキストベース名、そのテキストベース実体を格納するテキストファイル名、そのテキストベース3についての関連情報を格納する関連ファイル名、使用すべき検索アルゴリズム情報を管理するテキストベース管理機能。

【0016】（3c）テキストベース3の検索を新たに始める際に、テキストベース3のサイズ情報と上記（2c）のテキストベース管理機能で管理されている検索アルゴリズム情報と上記（1c）の検索サーバ管理機能で管理されている各検索サーバ1の使用可能メモリ量とに基づいてテキストベース3の分割数を決定する分割数決定機能。

【0017】（4c）上記（3c）の分割数決定機能により決定された分割数と前記の各検索サーバ1の使用可能メモリ量とに応じて、テキストベース3とその関連情報とを、各検索サーバ1のうち使用可能メモリ量の大きい検索サーバ1から順に割り当てる検索サーバ割当機能。

（5c）上記（4c）の検索サーバ割当機能による割当結果に基づいて、テキストベース3の分割部分31とその関連情報とを割当先の検索サーバ1へロードするためのローディング指示を、上記（2c）のテキストベース管理機能で管理されているテキストファイル名、関連情報ファイル名および検索アルゴリズム情報とともに割当先の各検索サーバ1へ送るローディング指示機能。

【0018】一方、この場合、各検索サーバ1は、以下のような各機能を有して構成される（請求項5）。

（1d）検索管理サーバ2からのローディング指示を受けると、このローディング指示に伴うテキストファイル名および関連情報ファイル名によってそれぞれ指定されるテキストベース3の分割部分31およびその関連情報

を読み込むファイルアクセス機能。

【0019】（2d）複数種類の検索アルゴリズムをそなえ、検索要求を受けると、検索管理サーバ2からの検索アルゴリズム情報により指定される検索アルゴリズムを用いて、上記（1d）のファイルアクセス機能により読み込んだテキストベース3の分割部分31に対する情報検索を行なう検索機能。

（3d）上記（2d）の検索機能による検索結果を検索要求元に返す検索結果送出機能。

10 【0020】また、図1に示す情報検索システムは、ユーザプロセスからの検索要求を受け付けるユーザ管理サーバをそなえ、このユーザ管理サーバが、ユーザプロセスからの検索要求に応じて、検索対象テキストベース3に対する検索を行なっている検索サーバ群情報の照会を検索管理サーバ2に行ない、その照会結果に基づいて実際の検索要求を各検索サーバ1に対して直接的に行なうように構成してもよい（請求項6）。

【0021】そして、この場合、上述のユーザ管理サーバは、以下のような各機能を有して構成される（請求項7）。

（1e）ユーザプロセスからの検索要求を受け付ける検索要求受付機能。

（2e）上記（1e）の検索要求受付機能により受け付けた検索要求に応じて、検索要求の検索対象テキストベース3に対する検索処理を割り当てられている検索サーバ群情報を検索管理サーバ2に照会する検索サーバ照会機能。

【0022】（3e）上記（2e）の検索サーバ照会機能により検索管理サーバ2から得られた検索サーバ群情報に基づいて、検索要求の検索対象テキストベース3に対する検索処理を割り当てられている各検索サーバ1に対して直接的に検索要求を送出する検索要求機能。

（4e）上記（3e）の検索要求機能の検索要求に対する各検索サーバ1からの応答をマージし検索結果として検索要求元のユーザプロセスに転送する検索結果転送機能。

【0023】また、この場合、検索管理サーバ2は、以下のような各機能を有して構成される（請求項8）。

（1f）検索サーバ1毎に、検索を担当しているテキストベース名とそのテキストベース3について検索を担当している分割部分31のデータ量とを含む検索サーバ情報を管理する検索サーバ管理機能。

【0024】（2f）検索対象テキストベース3のローディング先である検索サーバ群に関するローディング情報を管理するローディング情報管理機能。

（3f）検索対象テキストベース3毎に、そのテキストベース名、そのテキストベース実体を格納するテキストファイル名、そのテキストベース3についての関連情報を格納する関連ファイル名を管理するテキストベース管理機能。

【0025】(4f) ユーザ管理サーバから受けた照会要求の検索対象テキストベース3が未だロードされていない場合、そのテキストベース3のサイズ情報と上記

(1f) の検索サーバ管理機能で管理されている検索サーバ情報から得られる各検索サーバ1の負荷状況とに基づいてテキストベース3の分割数を決定する分割数決定機能。

【0026】(5f) 上記(4f)の分割数決定機能により決定された分割数と前記の各検索サーバ1の負荷状況とに応じて、テキストベース3とその関連情報とを、負荷の軽い検索サーバ1から順に割り当てる検索サーバ割当機能。

(6f) 上記(5f)の検索サーバ割当機能による割当結果に基づいて、テキストベース3の分割部分31とその関連情報とを割当先の検索サーバ1へロードするためのローディング指示を、上記(3f)のテキストベース管理機能により得られるテキストファイル名および関連情報ファイル名とともに割当先の各検索サーバ1へ送るローディング指示機能。

【0027】(7f) 上記(5f)の検索サーバ割当機能によりテキストベース3を割り当てられた検索サーバ群に関するローディング情報を、検索サーバ群情報としてユーザ管理サーバへ転送するローディング情報転送機能。なお、このとき上記(4f)の分割数決定機能により決定された分割数分の検索サーバ1を上記(5f)の検索サーバ割当機能により確保できない場合には、上記(4f)の分割数決定機能により決定されたテキストベース3の分割数を見直し、その分割数に応じて上記(5f)の検索サーバ割当機能、上記(6f)のローディング指示機能および上記(7f)のローディング情報転送機能による処理が行なわれる(請求項9)。

【0028】また、上述のユーザ管理サーバから受けた照会要求の検索対象テキストベース3が既にロードされている場合には、上記(7f)のローディング情報転送機能によって、上記(2f)のローディング情報管理機能で管理されている検索対象テキストベース3についてのローディング情報が、検索サーバ群情報としてユーザ管理サーバへ転送される(請求項10)。

【0029】さらに、上述のようにユーザ管理サーバから受けた照会要求の検索対象テキストベース3が既にロードされている場合、検索管理サーバ2では、次のような処理を行なうようにしてもよい(請求項11)。すなわち、上記(4f)の分割数決定機能により、テキストベース3のサイズ情報と上記(1f)の検索サーバ管理機能で管理されている検索サーバ情報から得られる各検索サーバ1の負荷状況とに基づいてテキストベース3の分割数を決定し、各検索サーバ1の中から所定値よりも負荷の軽い検索サーバ1を上記(4f)の分割数決定機能により決定された分割数だけ確保できる場合、上記

(5f)の検索サーバ割当機能により、前記所定値より

も負荷の軽い検索サーバ1を分割数だけ選択してテキストベース3とその関連情報とを割り当てる。

【0030】そして、前記(5a)のローディング指示機能により、上記(5f)の検索サーバ割当機能による割当結果に基づいて、テキストベース3の分割部分(A0~An)とその関連情報とを割当先の検索サーバ1へロードするためのローディング指示を、上記(3f)のテキストベース管理機能により得られるテキストファイル名および関連情報ファイル名とともに割当先の各検索サーバ1へ送り、上記(7f)のローディング情報転送機能によって、上記(5f)の検索サーバ割当機能によりテキストベース3を割り当てられた検索サーバ群に関するローディング情報を、検索サーバ群情報としてユーザ管理サーバへ転送する。

【0031】一方、各検索サーバ1の中から前記所定値よりも負荷の軽い検索サーバ1を上記(4f)の分割数決定機能により決定された分割数だけ確保できない場合は、上記(7f)のローディング情報転送機能により、上記(2f)のローディング情報管理機能で管理されている検索対象テキストベース3についてのローディング情報を、検索サーバ群情報としてユーザ管理サーバへ転送する。

【0032】このため、この場合の各検索サーバ1は、以下のような各機能を有して構成される(請求項12)。

(1g) 検索管理サーバ2からのローディング指示を受けると、ローディング指示に伴うテキストファイル名および関連情報ファイル名によってそれぞれ指定されるテキストベース3の分割部分31およびその関連情報を読み込むファイルアクセス機能。

【0033】(2g) ユーザ管理サーバから検索要求を受けると、上記(1g)のファイルアクセス機能により読み込んだテキストベース3の分割部分31に対する情報検索を行なう検索機能。

(3g) 上記(2g)の検索機能による検索結果をユーザ管理サーバに返す検索結果送出機能。

【0034】また、上述の検索管理サーバ2は、以下のような各機能を有して構成してもよい(請求項13)。

(1h) 検索サーバ1毎に、検索を担当しているテキストベース名、そのテキストベース3について検索を担当している分割部分31のデータ量、使用可能メモリ量を含む検索サーバ情報を管理する検索サーバ管理機能。

【0035】(2h) 検索対象テキストベース3のローディング先である検索サーバ群に関するローディング情報を管理するローディング情報管理機能。

(3h) 検索対象テキストベース3毎に、そのテキストベース名、そのテキストベース実体を格納するテキストファイル名、そのテキストベース3についての関連情報を格納する関連ファイル名、使用すべき検索アルゴリズム情報を管理するテキストベース管理機能。

【0036】(4h) ユーザ管理サーバから受けた照会要求の検索対象テキストベース3が未だロードされていない場合、テキストベース3のサイズ情報と上記(3h)のテキストベース管理機能で管理されている検索アルゴリズム情報と上記(1h)の検索サーバ管理機能で管理されている各検索サーバ1の使用可能メモリ量とに基づいてテキストベース3の分割数を決定する分割数決定機能。

【0037】(5h) 上記(4h)の分割数決定機能により決定された分割数と各検索サーバ1の使用可能メモリ量とに応じて、テキストベース3とその関連情報とを、使用可能メモリ量の大きい検索サーバ1から順に割り当てる検索サーバ割当機能。

(6h) 上記(5h)の検索サーバ割当機能による割当結果に基づいて、テキストベース3の分割部分31とその関連情報とを割当先の検索サーバ1へロードするためのローディング指示を、上記(3h)のテキストベース管理機能により得られるテキストファイル名および関連情報ファイル名とともに割当先の各検索サーバ1へ送るローディング指示機能。

【0038】(7h) 上記(5h)の検索サーバ割当機能によりテキストベース3を割り当てられた検索サーバ群に関するローディング情報を、検索サーバ群情報としてユーザ管理サーバへ転送するローディング情報転送機能。なお、この場合も、上記(4h)の分割数決定機能により決定された分割数分の検索サーバ1を上記(5h)の検索サーバ割当機能により確保できない場合には、上記(4h)の分割数決定機能により決定されたテキストベース3の分割数を見直し、その分割数に応じて上記(5h)の検索サーバ割当機能、上記(6h)のローディング指示機能および上記(7h)のローディング情報転送機能による処理が行なわれる(請求項14)。

【0039】また、ユーザ管理サーバから受けた照会要求の検索対象テキストベースが既にロードされている場合は、上記(7h)のローディング情報転送機能により、上記(2h)のローディング情報管理機能で管理されている検索対象テキストベース3についてのローディング情報が、検索サーバ群情報としてユーザ管理サーバへ転送される(請求項15)。

【0040】なお、上述のようにユーザ管理サーバから受けた照会要求の検索対象テキストベース3が既にロードされている場合、この検索管理サーバ2では、次のような処理を行なうようにしてもよい(請求項16)。すなわち、上記(4h)の分割数決定機能により、テキストベース3のサイズ情報と上記(3h)のテキストベース管理機能で管理されている検索アルゴリズム情報と上記(1h)の検索サーバ管理機能で管理されている各検索サーバ1の使用可能メモリ量とに基づいてテキストベース3の分割数を決定する。

【0041】そして、各検索サーバ1の中から所定値よ

りも使用可能メモリ量の大きい検索サーバ1を上記(4h)の分割数決定機能により決定された分割数だけ確保できる場合、上記(1h)の検索サーバ割当機能が、前記所定値よりも使用可能メモリ量の大きい検索サーバ1を分割数だけ選択してテキストベース3とその関連情報とを割り当て、上記(6h)のローディング指示機能によって、上記(1h)の検索サーバ割当機能による割当結果に基づいて、テキストベース3の分割部分31とその関連情報とを割当先の検索サーバ1へロードするためのローディング指示を、上記(3h)のテキストベース管理機能により得られるテキストファイル名、関連情報ファイル名および検索アルゴリズム情報とともに割当先の各検索サーバ1へ送り、上記(7h)のローディング情報転送機能が、上記(1h)の検索サーバ割当機能によりテキストベース3を割り当てられた検索サーバ群に関するローディング情報を、検索サーバ群情報としてユーザ管理サーバへ転送する。

【0042】一方、複数の検索サーバ1の中から前記所定値よりも使用可能メモリ量の大きい検索サーバ1を上記(4h)の分割数決定機能により決定された分割数だけ確保できない場合、上記(7h)のローディング情報転送機能により、上記(2h)のローディング情報管理機能で管理されている検索対象テキストベース3についてのローディング情報を、検索サーバ群情報としてユーザ管理サーバへ転送する。

【0043】なお、この場合、検索サーバ1は、以下のような各機能を有して構成される(請求項17)。

(1i) 検索管理サーバ2からのローディング指示を受けると、ローディング指示に伴うテキストファイル名および関連情報ファイル名によってそれぞれ指定されるテキストベース3の分割部分31およびその関連情報を読み込むファイルアクセス機能。

【0044】(2i) 複数種類の検索アルゴリズムをそなえ、ユーザ管理サーバから検索要求を受けると、検索管理サーバ2からの検索アルゴリズム情報により指定される検索アルゴリズムを用いて、上記(1i)のファイルアクセス機能により読み込んだテキストベース3の分割部分31に対する情報検索を行なう検索機能。

(3i) 上記(2i)の検索機能による検索結果をユーザ管理サーバに返す検索結果送出機能。

【0045】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態を説明する。

(a) 情報検索システム全体の構成説明

図2は本発明の一実施形態としての情報検索システムの構成を示すブロック図で、この図2において、4はユーザ管理サーバ、5は検索管理サーバ、6は検索サーバ、7は複数のテキストベース(TB)8を保持するファイルシステム(データベース)である。

【0046】なお、この図2に示す情報検索システム

は、各サーバ（ユーザ管理サーバ4、検索管理サーバ5、検索サーバ6）が、それぞれ複数のPE（Processor Element）として分散構築されていることを想定しているが、これらの各サーバのうちいくつかのサーバが同一のPEとして構築される場合もある。また、各サーバは、それぞれこの図2に示す数以上存在してもよい。

【0047】ここで、各ユーザ管理サーバ4は、システムの利用者（ユーザ）を統括して管理し、ファイルシステム7に保持されているテキストベース8に対するユーザからの検索要求に応じて、各検索サーバ6に検索要求を出すものであるが、本実施形態では、それぞれWAIS（Wide Area Information Server）プロトコルなどの通信プロトコルに従ってクライアントプロセス（ユーザプロセス）から送信されてくる検索要求を受け付け、受け付けた検索要求に応じて、テキストベース8に対する検索を行なっている検索サーバ（群）6の情報を、後述する検索管理サーバ5に照会して獲得し、獲得した照会結果に基づいて実際の検索要求を各検索サーバ6に対して直接的に行なうように構成されている。

【0048】また、検索管理サーバ5は、各検索サーバ6の動作を管理するためのもので、ここでは、ファイルシステム7に保持されている検索対象のテキストベース8と、そのテキストベース8についての関連データ（後述するインデックス情報など）とをそれぞれ分割し、その分割部分（分割テキストベース）81とそれぞれに対応する関連データとを組にして各検索サーバ6のうちの全部または全部に割り当てるように構成されている。

【0049】例えば、図2では、この検索管理サーバ5により、テキストベース8とそのテキストベース8についての関連データとがそれぞれ分割され、その分割部分81とこれらの分割部分81に対応する関連データとが各検索サーバ6に割り当てられることを表している。さらに、各検索サーバ6は、それぞれ、テキストベース8に対する情報検索を実際に行なうもので、ここでは、図9により後述するような複数種類の検索エンジン（検索アルゴリズム）61-1～61-m（mは自然数）をそなえ、上述の検索管理サーバ5によって割り当てられた分割部分81に対する情報検索が、検索管理サーバ5により指定される検索エンジン61-1～61-mのいずれかを用いて、並列かつ独立に行なわれるように構成されている。

【0050】図3は上述の各ユーザ管理サーバ4の詳細構成を示すブロック図で、この図3に示すように、各ユーザ管理サーバ4は、それぞれ、システムを利用するユーザ毎に作成されるユーザ固有情報テーブル41、外部命令プロトコル送受信機能42、ユーザ管理サーバ命令解析部43、ユーザ固有情報登録・更新・照会機能44、検索サーバ照会機能45、検索要求機能46及び検索マッチデータ回復機能47を有して構成される。

【0051】ここで、各ユーザ固有情報テーブル41

は、それぞれ、ユーザ名、アクセス（検索対象）テキストベース名、アクセス識別子（AD：Access Descriptor）などの情報をユーザ毎に管理するためのもので、例えば図4に示すように、ユーザ名（User name）41A、ユーザ端末の収容位置などを示すユーザアドレス（User address）41B、ADポインタ（*AD pointer）41C、マッチポイントテーブルポインタ（*Match point tp）41Dなどの各種の固有情報を格納するリスト構造となっている。

10 【0052】なお、ADポインタ41Cは、検索対象のテキストベース8毎に作成される検索サーバリスト411の先頭アドレスを指すポインタで、新たに作成される検索サーバリスト411の先頭アドレスを既に作成された検索サーバリスト411の次リスト参照用ポインタ（*next ptr）411aに順次格納してリンクさせてゆくことにより、このADポインタ41Cから全ての検索サーバリスト411を参照できるように管理されている。

20 【0053】また、マッチポイントテーブルポインタ41Dは、検索条件（検索すべきキーワードなど）毎に作成される検索条件リスト412の先頭アドレスを指すポインタで、新たに作成される検索条件リスト412の先頭アドレスを既に作成された検索条件リスト412の次リスト参照用ポインタ（*next ptr）412aに順次格納してリンクさせてゆくとともに、検索条件にマッチしたデータのアドレス情報を格納するマッチポイントテーブル413の先頭アドレスをそれぞれ対応する検索条件リスト412のテーブルポインタ（*table ptr）に格納してリンクさせることにより、このマッチポイントテーブルポインタ41Dから全ての検索条件リスト412、マッチポイントテーブル413を参照できるように管理されている。

30 【0054】一方、図3において、外部命令パケット送受信機能（検索要求受付機能）42は、クライアント（ユーザ）プロセスなどの外部との通信の際に遣り取りされる命令パケットを送受信して、ユーザプロセスから命令パケットとして送信されてくる検索要求を受け付けるものであり、ユーザ管理サーバ命令解析部43は、この外部命令パケット送受信機能42を通じて遣り取りされる命令パケットから検索要求などの各種の命令を解析するものである。

40 【0055】さらに、ユーザ固有情報登録・更新・照会機能44は、このユーザ管理サーバ命令解析部41での解析結果に応じて、ユーザ固有情報テーブル41の作成、ユーザ固有情報テーブル41内の各種情報の登録、更新、照会などを行なうものである。また、検索サーバ照会機能45は、外部命令プロトコル送受信機能42、ユーザ管理サーバ命令解析部43を通じて受け付けられたユーザからの検索要求に応じて、検索要求されたテキストベース（群）8に対する検索処理を割り当てられて
50 いる検索サーバ群情報（検索サーバ6のID群など）を

検索管理サーバ5に照会するもので、この検索サーバ照会機能43により、ユーザの指定した検索対象のテキストベース8がどの検索サーバ6に保持されているかという情報（検索サーバ群情報）を獲得できるようになっている。

【0056】さらに、検索要求機能46は、この検索サーバ照会機能45により検索管理サーバ5から得られた検索サーバ群情報に基づいて、検索要求の出ているテキストベース8に対する検索処理を割り当てられている各検索サーバ6に対して直接的に検索要求を送出するものであり、この検索要求機能46の検索要求に応じた各検索サーバ6による検索結果（各検索サーバ6でのアドレス情報）は、マッチポイントデータとしてマッチポイントテーブル413に格納されるようになっている。

【0057】検索マッチデータ回復機能47は、外部命令パケット送受信機能42、ユーザ管理サーバ命令解析部43を通じて受け付けられたユーザからの検索マッチデータ回復要求に応じて、各検索サーバ6から読み出されたデータ（ユーザ固有情報テーブル41内のマッチポイントテーブル413のマッチポイントデータに対応するデータ）をマージして、外部命令パケット送受信機能42を通じて検索要求元のユーザプロセスに転送するものである。

【0058】次に、図5は上述の検索管理サーバ5の詳細構成を示すブロック図で、この図5に示すように、検索管理サーバ5は、テキストベース8毎に作成されるテキストベース（TB）管理テーブル51、アクセス識別子（AD）管理テーブル52、各検索サーバ6毎に作成される検索サーバ情報テーブル53、検索管理サーバ命令解析部54、テーブル更新機能55、アクセス識別子（AD）割当機能56、データ分割数決定機能57、検索サーバ割当機能58及びローディング指示機能59を有して構成され、基本的に、この検索管理サーバ5は、各テーブル51～53に基づいて動作するようになっている。

【0059】ここで、各TB管理テーブル（テキストベース管理機能）51は、それぞれ、複数のテキストベース8毎に、テキストベース名とその原データ及び関連データ（インデックス情報など）を管理するもので、後述するように各検索サーバ6に対してテキストベース8の分割部分81のローディング指示を出す際に使用される。

【0060】具体的に、各TB管理テーブル51には、例えば図6に示すように、それぞれテキストベース名、そのテキストベース8の分割部分81を保持するテキストファイル名（1～n：nはテキストベース8の分割数）、そのサイズ情報（1～n）、アクセスユーザ数、検索サーバID群、そのテキストベース8についてのインデックスなどの関連データを格納する関連ファイル名（1～n）、そのサイズ情報（1～n）、テキストベ

ス8毎に使用すべき検索エンジン（検索アルゴリズム）61-1～61-mなどを格納するリスト構造になっている。

【0061】なお、アクセスユーザ数は現在そのテキストベース8を用いて検索を行なっている総ユーザ数、検索サーバID群はこのテキストベースの検索を担当している検索サーバ6のID群、使用可能メモリ量はテキストベース8の割当／解放に伴って増減する検索サーバ6の残りメモリ量である。一方、AD管理テーブル（ローディング情報管理機能）52は、例えば図7に示すように、AD毎に、対応する参照ユーザ数、テキストベース名、検索対象テキストベース8をローディング（保持）している検索サーバ（群）6に関するローディング情報（検索サーバ6のID群など）を管理するものであり、検索サーバ情報テーブル（検索サーバ管理機能）53は、例えば図8に示すように、検索サーバ6のID、検索ユーザ数、総検索テキストベース量、使用可能メモリ量、検索を担当しているテキストベース名、そのテキストベース8について検索を担当している分割部分81のデータ量などを含む検索サーバ情報を管理するもので、これらの各種情報に基づいて、各検索サーバ6の残りメモリ量（又は負荷状況）などを判断できるようになっている。

【0062】なお、検索ユーザ数はその検索サーバ6を用いて検索を行なっているユーザ数、総検索テキストベース量はその検索サーバ6で現在検索を担当しているテキストベース量の合計である。さらに、図5において、検索管理サーバ命令解析部54は、外部との通信の際に遣り取りされる命令パケットから各種の命令を解析するものであり、テーブル更新機能55は、上述のTB管理テーブル51、AD管理テーブル52及び検索サーバ情報テーブル53の各種情報に対する更新などを行なうものであり、AD割当機能56は、TB管理テーブル51、検索サーバ情報テーブル53で管理されている各種の情報に基づいて、AD管理テーブル52のADに対応する情報の割当／獲得などを行なうもので、例えば、ユーザ管理サーバ4から新たなテキストベース8に対する照会要求を受けると、新たなADを割り当てるとともに、AD管理テーブル52にそのADについての領域を獲得して各種パラメータ（図7参照）を設定するものである。

【0063】また、データ分割数決定機能57は、ユーザ管理サーバ4から受けた照会要求の検索対象テキストベース8が未だロードされていない場合、そのテキストベース8のサイズ情報と、TB管理テーブル51で管理されている使用検索アルゴリズム情報と、検索サーバ情報テーブル53で管理されている各検索サーバ6の使用可能メモリ量とに基づいてテキストベース8の分割数を決定するもので、検索サーバ割当機能58は、このデータ分割数決定機能57により決定された分割数と各検

索サーバ6の使用可能メモリ量(又は、負荷状況)とに応じて、テキストベース8とその関連データとを、使用可能メモリ量の大きい検索サーバ6(又は、負荷状況の軽い検索サーバ6)から順に割り当てるものである。

【0064】さらに、ローディング指示機能59は、この検索サーバ割当機能58による割当結果に基づいて、テキストベース8の分割部分とその関連データとを割当先の検索サーバ6へロードするためのローディング指示を、TB管理テーブル51により得られるテキストファイル名および関連ファイル名とともに割当先の各検索サーバ6へ送るものであるが、本実施形態では、検索対象テキストベース8の分割部分81を割り当てられた検索サーバ(群)6に関するローディング情報を、実際に分割部分81に対する検索を行なう検索サーバ群情報としてユーザ管理サーバ4へ転送するローディング情報転送機能も兼ねている。

【0065】つまり、この検索管理サーバ5は、例えば、ユーザ管理サーバ4から照会要求のあったテキストベース(ユーザが検索要求した検索対象テキストベース)8がその時点で未だ検索サーバ6にロードされていない場合は、そのテキストベース8の大きさ、使用検索アルゴリズム情報、各検索サーバ6の使用可能メモリ量(又は、負荷状況)などに応じて、テキストベース8の分割数を決定し、この分割数に基づいて、各検索サーバ6の中から必要な数の検索サーバ6を選択して、各検索サーバ6に、それぞれが担当する部分(検索対象テキストベース8の分割部分81)のみをロードさせたのち、そのローディング情報(検索を割り当てられた検索サーバ6のIDなど)を照会元のユーザ管理サーバ4に返すようになっているのである。

【0066】なお、この検索管理サーバ5は、上述のようにデータ分割数決定機能57により決定された分割数分の検索サーバ6を、検索サーバ割当機能58により確保できない場合には、後述するように、データ分割数決定機能57により決定されたテキストベース8の分割数を見直し、見直された分割数に応じて、再度、検索サーバ割当機能58、ローディング指示機能59による処理が行なわれるようになっている。

【0067】また、本実施形態では、ユーザの検索要求した検索対象テキストベース8が検索サーバ6に既にロードされている場合は、ローディング指示機能59により、AD管理テーブル52に既に登録され管理されている検索対象テキストベース8についてのローディング情報が検索サーバ群情報としてユーザ管理サーバ4へ転送されるようになっているが、検索対象テキストベース8が既にロードされている場合は、後述するように、所定値よりも使用可能メモリ量の大きい(又は、所定値よりも負荷の軽い)検索サーバ6を確保できれば、確保した他の検索サーバ6にも検索対象テキストベース8を割り当てるようにしてもよい。

【0068】次に、図9は上述の各検索サーバ6の詳細構成を示すブロック図で、この図9に示すように、各検索サーバ6は、それぞれ、図2により前述したように複数種類の検索エンジン61-1~61-mをそなえるほか、検索サーバ命令解析部62、ファイルアクセス機能63、AD毎に作成されるテキストベース/インデックス(TB/Index)管理テーブル64、各検索サーバ6の存在するPEのメモリ上にキャッシュされるテキスト領域65、検索機能66、検索マッチ情報格納機能67、マッチポイントテーブル68、検索結果送出機能69及び原データ転送機能70を有して構成される。

【0069】ここで、各検索エンジン61-1~61-mは、それぞれ、情報検索の際に検索精度を重視するか検索速度を重視するかなどの条件に応じた検索方法(検索アルゴリズム)による検索を実行するもので、後述する検索機能66により、情報検索の際に使用すべき検索エンジン61-1~61-mが選択されて用いられるようになっている。

【0070】また、検索サーバ命令解析部62は、ユーザ管理サーバ4から受信する命令パケットから行なうべき処理を解析するものであり、ファイルアクセス機能63は、検索管理サーバ5からのローディング指示を受けると、そのローディング指示に伴うテキストファイル名および関連ファイル名によってそれぞれ指定されるテキストファイル(テキストベース8の分割部分81の実体を格納したファイル)71'及びそのテキストファイル71'の関連データ(インデックス情報など)が格納された関連ファイル72'をテキスト領域65に読み込むものである。

【0071】さらに、各TB/Index管理テーブル64は、それぞれ、上述のごとくファイルアクセス機能63によりテキスト領域65に格納されたテキストファイル71'と関連ファイル72'とのテキスト領域65におけるアドレス情報や、検索サーバ6の使用可能メモリ量(残りメモリ容量)、情報検索時に使用すべき検索エンジン61-1~61-mなどの情報を管理するものである。

【0072】例えば、このTB/Index管理テーブル64には、それぞれ、図10に示すように、テキストベースポインタ(*TB ptr)641、関連データポインタ(*con ptr)642を始め、テキストベース(TB)ファイル名、オープンしたテキストベースのファイル識別子(TB用file descriptor)、関連ファイル名、オープンした関連ファイルのファイル識別子(関連ファイル用fd)、使用可能メモリ量、使用検索エンジンなどの各種情報がAD(AD#0, AD#1, ...)毎に格納されるようになっている。

【0073】そして、この場合も、例えば、テキストベースポインタ641に、読み込んだテキストファイル71'の先頭アドレス(テキスト領域65内の格納位置を

示すアドレス情報)を格納し、関連データポインタ642に、読み込んだ関連ファイル72'の先頭アドレス情報をそれぞれ格納することによって、各検索サーバ65で自己が保持する全てのテキストファイル71'と関連ファイル72'についての情報を参照できるように管理される。

【0074】また、検索機能66は、検索サーバ命令解析部62を通じてユーザ管理サーバ4から検索要求を受けると、検索エンジン61-1~61-mのうち、検索管理サーバ5からの検索アルゴリズム情報により指定される検索エンジンを用いて、ファイルアクセス機能63によりテキスト領域65に読み込んだテキストベース8の分割部分81(テキストファイル71', 関連ファイル72')に対する情報検索を行なうものである。

【0075】さらに、検索マッチ情報格納機能67は、この検索機能66による検索結果をマッチポイントテーブル68に一旦格納させるもので、例えば、この場合は、キーワードなどの検索条件にマッチしたデータのテキスト領域65における位置(アドレス情報)がマッチポイントデータとしてマッチポイントテーブル68に格納されるようになっている。

【0076】また、検索結果送出機能69は、検索マッチ情報格納機能67により、マッチポイントテーブル68に格納されたマッチポイントデータ(アドレス情報)を、検索対象テキストベース8に対する検索結果として検索要求元のユーザ管理サーバ4に返すものであり、原*

*データ転送機能70は、ユーザ管理サーバ4からのマッチデータ回復要求に応じて、このマッチデータ回復要求とともに送られてくるマッチポイントデータに対応するテキスト領域65内の原データを取り出してユーザ管理サーバ4へ転送するものである。

【0077】つまり、各検索サーバ6は、それぞれ、ファイルアクセス機能63により各自が検索担当するテキストベース8の分割部分81のみをテキスト領域65に読み込んで、検索機能66によりその分割部分81に対する情報検索を行なうことによって、検索対象テキストベース8の分割部分81に対する情報検索を並列かつ独立に行なえるようになっている。

【0078】なお、以下に示す表1~表3はそれぞれ上述のユーザ管理サーバ4、検索管理サーバ5、検索サーバ6間の通信時に付加されるパラメータ(通信プロトコル)の一例を示す表で、例えば、表1はユーザ管理サーバ4と検索管理サーバ5との間での通信プロトコルを示し、ユーザ管理サーバ4から検索管理サーバ5にテキストベース照会要求を出す場合は、パラメータとしてそのテキストベース名が付加され、この照会要求に対する検索管理サーバ5からの応答に、パラメータとしてAD、その照会要求のあったテキストベース8を保持している検索サーバ(群)6のID(群)が付加されることを表している。

【0079】

【表1】

ユーザ管理サーバと検索管理サーバとの間の通信プロトコル

	ユーザ管理サーバ ←検索管理サーバ (付加されるパラメータ)	検索管理サーバ ←ユーザ管理サーバ (付加されるパラメータ)
テキストベース照会要求時	・テキストベース名	・アクセス識別子(AD) ・検索サーバID(群)
検索結果返却時	・アクセス識別子(AD)	なし

【0080】

※ ※ 【表2】

検索管理サーバと検索サーバとの間の通信プロトコル

	検索管理サーバ←検索サーバ (付加されるパラメータ)	検索サーバ←検索管理サーバ (付加されるパラメータ)
テキストベースローディング指示時	・テキストファイル名 ・関連ファイル名 ・アクセス識別子(AD) ・テキストベースの分割数	なし
テキストベース戻却指示時	・アクセス識別子(AD)	なし
検索サーバ状態通知指示時	なし	・担当テキストベース量 ・総担当テキストベース量 ・使用可能メモリ量

【0081】

【表3】

ユーザ管理サーバと検索サーバとの間の通信プロトコル

	ユーザ管理サーバ→検索サーバ (付加されるパラメータ)	検索サーバ→ユーザ管理サーバ (返送される情報)
検索要求時	・検索キーワード ・アクセス識別子 (AD) ・ユーザ名 (User name)	・ヒット数 ・検索キーワードのTB上 でのアドレス (返数) ・ユーザ名 (User name)
戻データ返送 要求時	・アクセス識別子 (AD) ・戻データアドレス (返数) ・1アイテム当たりの 返送サイズ ・ユーザ名 (User name)	・返送テキストベース群 ・ユーザ名 (User name)

【0082】以下、上述のごとく構成された情報検索システムの動作について、ユーザ管理サーバ4、検索管理サーバ5、検索サーバ6ごとに詳述する。

(b) ユーザ管理サーバ4の動作説明

まず、ユーザ管理サーバ4の動作について、図11～図14に示す各フローチャート(ステップS1～S16)を用いて説明する。

【0083】例えば、ユーザプロセスからユーザ管理サーバ4に対して何らかの処理要求があったとする。すると、ユーザ管理サーバ4では、図11に示すように、受け付けた要求が、検索対象テキストベース8のオープン処理、検索処理、検索マッチデータ回復処理(検索結果の転送処理)のいずれであるかがユーザ管理サーバ命令解析部43(図3参照)で判定される(ステップS1～S3)。

【0084】この判定の結果、ユーザプロセスからの要求が、例えばテキストベース8のオープン処理であった場合(ステップS1でYESと判定された場合)は、図12に示すように、要求のあったユーザに対応するユーザ固有情報テーブル41(図4参照)が既に作成されているかどうか判定され(ステップS4)、作成されていない場合(ステップS4でNOと判定された場合)は、ユーザ固有情報登録・更新・照会機能44により、そのユーザに対応するユーザ固有情報テーブル41を新たに作成して、要求元のユーザ名、ユーザアドレス、検索するテキストベース名などを設定したのち(ステップS5)、検索サーバ照会機能45により、ユーザから指定された検索対象テキストベース8のテキストベース名を付加したテキストベース(検索サーバ)照会要求を検索管理サーバ5に送出する(ステップS6)。

【0085】なお、要求のあったユーザに対応するユーザ固有情報テーブル41が既に作成されていた場合は、そのユーザ固有情報テーブル41に基づいて、同様に、テキストベース(検索サーバ)照会要求を検索管理サーバ5に送出する(ステップS4のYESルートからステップS6)。その後、ユーザ管理サーバ4は、検索管理サーバ5からの応答を受信するまで処理を待機し(ステップS7のNOルート)、検索管理サーバ5からの応答を受信すると、ユーザ固有情報登録・更新・照会機能4

4によって、ユーザ固有情報テーブル41のADポイント41Cが指すリスト領域に新しい検索条件リスト412を作成し、検索管理サーバ5からの応答に付加されているパラメータ(AD、検索サーバID群、テキストベース名など)をそのリスト412に書き込んで登録する(ステップS7のYESルートからステップS8)。

【0086】以上の処理が完了すると、ユーザ管理サーバ4は、要求のあったユーザに対してテキストベース8のオープン処理が完了した旨を通知したのち、次の要求に対する待受動作状態となる(ステップS9)。次に、ユーザプロセスから検索処理が要求され、ユーザ管理サーバ命令解析部43で受け付けた要求が検索処理であると判定されると(図11のステップS2でYESと判定されると)、図13に示すように、ユーザ管理サーバ4では、まず、要求元のユーザ名(User name)に該当するユーザ固有情報テーブル41をユーザ固有情報登録・更新・照会機能44により特定し、その項目(リスト411)からユーザが指定したテキストベース名に対応するADを抽出したのち、検索要求機能46により、このADとともにリスト411で指定されている検索サーバのID(群)に対応する検索サーバ(例えば、検索サーバ6)に検索要求(検索コマンド)を送出する(ステップS10)。

【0087】つまり、ユーザ管理サーバ4は、ユーザプロセスからの検索要求に応じて、検索対象テキストベース8に対する検索を行なっている検索サーバ群情報(検索サーバ6のID群など)の照会を検索管理サーバ5に行ない、その照会結果に基づいて、実際の検索要求を各検索サーバ6に直接的に行なう。なお、ユーザにより検索対象のテキストベース名が指定されていなかった場合は、そのユーザが保持している全てのADに対応するテキストベース8が検索対象となる。

【0088】そして、ユーザ管理サーバ4は、検索要求を送出した全ての検索サーバ6からの応答を受信するまで処理を待機し(ステップS11のNOルート)、全ての検索サーバ6からの応答を受信した時点で、ユーザ固有情報テーブル41にリンクされるマッチポイントテーブル413(図4参照)を必要サイズ分だけ割り当て、それぞれに受信したマッチポイントデータを書き込

んだのち(ステップS12)、要求元のユーザにマッチ数を返して、次の要求に対する待受動作状態となる(ステップS13)。

【0089】次に、その後、ユーザプロセスから検索マッチデータ回復処理(検索結果の転送処理)が要求され、その旨がユーザ管理サーバ命令解析部43で認識されると(図11のステップS3でYESと判定されると)、図14に示すように、ユーザ管理サーバ4は、まず、ユーザ固有情報登録・更新・照会機能44によって、要求元のユーザ名から該当するユーザ固有情報テーブル41を特定し、そのリスト412に格納されているマッチポイントテーブル413より、検索対象テキストベース8内でのアドレス(マッチポイントデータ)を抽出して検索サーバ6への転送要求メッセージを作成する。

【0090】その後、この転送要求メッセージは、検索サーバ群情報(検索サーバ6のID群)に基づいて、検索を行なった各検索サーバ6にADとともに送出される(ステップS14)。なお、ユーザにより転送要求のテキストベース名が指定されていなかった場合は、そのユーザが保持している全てのADに対応するマッチポイントテーブル413のマッチポイントデータが転送処理対象となる。

【0091】そして、ユーザ管理サーバ4は、転送要求を出した全ての検索サーバ6からの応答を受信するまで処理を待機し(ステップS15のNOルート)、全検索サーバ6からの応答を受信した時点で、AD毎に検索サーバ6から転送されてきたデータを検索マッチデータ回復機能47により編集(マージ)して、これを外部命令パケット送受信機能42により検索結果としてユーザに転送する(ステップS15のYESルートからステップS16)。

【0092】このように、上述のユーザ管理サーバ4では、ユーザプロセスからの検索要求に応じて、検索管理サーバ5に、検索対象のテキストベース8に対する情報検索を行なっている検索サーバ6に関する情報を照会することによって、常に、検索対象のテキストベース8を保持している検索サーバ6に実際の検索要求を直接行なうことができるので、ユーザプロセスからの検索要求に対して遅延なく、後述する各検索サーバ6に、検索対象テキストベース8に対する情報検索処理を実行させることができる。

【0093】(c) 検索管理サーバ5の動作説明

次に、以下では、検索管理サーバ5の動作について、図15～図17に示す各フローチャート(ステップS21～S39)を用いて説明する。まず、検索管理サーバ5は、図15に示すように、ユーザ管理サーバ4から検索テキストベースの照会要求があるまで待機状態となっている(ステップS21のNOルート)。この状態で、例えば、図12のステップS6により前述したように、ユ

ーザ管理サーバ4からテキストベース8についての照会要求が送出されると、検索管理サーバ5では、この照会要求を受けて、照会要求のあったテキストベース8が既にAD管理テーブル52に存在している(登録されている)かどうかAD割当機能56によって判定される(ステップS21のYESルートからステップS22)。

【0094】そして、照会要求のあったテキストベース8が既に登録され検索サーバ6にロードされている場合は、テーブル更新機能55によりAD管理テーブル52の参照ユーザ数(図7参照)を1つ増やしたのち(ステップS23)、既に登録されている対応するADと検索サーバ群情報(検索サーバ6のID群など)とが照会要求元のユーザ管理サーバ4に返される(ステップS24)。

【0095】つまり、この検索管理サーバ5は、照会要求のあった検索対象テキストベース8が既に検索サーバ6にロードされている場合は、検索対象テキストベース8を保持している検索サーバを新たに特定せずに、既にAD管理テーブル52に登録されている検索サーバ6に関する情報をユーザ管理サーバ4に転送するのである。従って、この場合、照会要求元のユーザ管理サーバ4が実際の検索要求を各検索サーバ6へ出すまでの処理時間が大幅に短縮される。

【0096】一方、照会要求のあったテキストベース8がロードされていない場合(ステップS22でNOと判定された場合)は、データ分割数決定機能57により、図16に示すように、TB管理テーブル51(図6参照)から該当するテキストベース8に関する情報を特定し、その実体を格納しているテキストファイル71'のサイズ情報、各TB管理テーブル51で管理されている使用検索アルゴリズム情報、各検索サーバ情報テーブル53で管理されている各検索サーバ6の使用可能メモリ量に基づいて、検索に必要な総メモリ量“a”，検索サーバ6が余裕をもって情報検索を実行できる1PE当たりの推奨メモリ量“b”，1PE当たりの残存メモリ量の最低ラインである最低メモリ量“c”が算出される(ステップS25、S26)。

【0097】さらに、データ分割数決定機能57では、上述のごとく算出された各メモリ量“a”，“b”，“c”を用いて、分割数“nPE”，最小分割数“nPE_{min}”が、それぞれ、 $(a+b-1)/b$ ， $(a+c-1)/c$ なる演算により決定されるとともに(ステップS27)、検索サーバ情報テーブル53(図8参照)に基づいて、使用可能メモリ量が推奨メモリ量“b”を超える検索サーバ数“d”が抽出され(ステップS28)、この検索サーバ数“d”が分割数“nPE”以上であるかどうか判定される(ステップS29)。

【0098】この結果、検索サーバ数“d”が分割数“nPE”以上であった場合(ステップS29でYES

と判定された場合)、つまり、上述のごとくデータ分割数決定機能57により決定された分割数“nPE”分の検索サーバ6を確保できる場合は、検索サーバ割当機能58により、使用可能メモリ量が推奨メモリ量“b”を超える“d”個分の検索サーバ6から“nPE”個分の検索サーバ6が抽出されるとともに、検索サーバ情報テーブル53における各種の情報が、テーブル更新機能55により以下のように更新される(ステップS29のYESルートからステップS30)。

【0099】・「検索ユーザ数」に1加算

・「総検索テキストベース量」に推奨メモリ量“b”を加算

・「使用可能メモリ量」から推奨メモリ量“b”を減算

・「検索テキストベース名」を登録

・「分割数」に“nPE”を登録

・「データ量」に推奨メモリ量“b”を登録

つまり、この検索管理サーバ5は、データ分割数決定機能57により決定された検索対象テキストベース8の分割数“nPE”と、各検索サーバ6の使用可能メモリ量とに応じて、検索サーバ割当機能58により、テキストベース8の実体を格納しているテキストファイル71'とその関連データを格納している関連ファイル72'とを、使用可能メモリ量の大きい検索サーバ6から順に割り当ててようになっているのである。

【0100】従って、この場合、特定の検索サーバ6の負荷が大きくなり過ぎるようなことはなく、各検索サーバ6は、常に、最適な負荷状況で高速に情報検索処理を行なうことができる。その後、検索管理サーバ5は、このような検索サーバ割当機能58による割当結果に基づいて、例えば、検索対象テキストベース8の分割部分81(テキストファイル71')と、関連ファイル72'とを割当先の各検索サーバ6へ分散してロードさせるためのローディング指示を、ローディング指示機能59により各検索サーバ6に送出する(ステップS31)。

【0101】そして、AD割当機能56により、TB管理テーブル51、検索サーバ情報テーブル53に基づいて、AD管理テーブル52に空き項目を1つ割り当て、参照ユーザ数に“1”、対応するテキストベース名に検索対象テキストベース8のファイル名、検索対象テキストベース8の分割部分81を保持している各検索サーバ6のID群を設定したのち(ステップS32)、対応するADと各検索サーバ6のID群を照会要求元のユーザ管理サーバ4に伝送して処理を終える(ステップS33)。

【0102】一方、検索サーバ数“d”が分割数“nPE”よりも小さかった場合(ステップS29でYESと判定された場合)、つまり、分割数“nPE”分の検索サーバを確保できない場合には、図17に示すように、改めて、データ分割数決定機能57により、使用可能メモリ量が最低メモリ量“c”を超える検索サーバ(推奨

メモリ量“b”分のメモリ量は残っていないが最低限必要なメモリ量“c”は残っている検索サーバ)の数

“e”を検索サーバ情報テーブル53から抽出して、分割数“nPE”を見直す(ステップS34)。

【0103】そして、上述のごとく抽出された検索サーバ数“e”は、図16のステップS27において算出された最小分割数“nPE_{min}”と比較され(ステップS35)、 $e < nPE_{min}$ であった場合(ステップS35でNOと判定された場合)、つまり、使用可能メモリ量(残りメモリ量)に多少余裕のある検索サーバ数“e”が最小分割数“nPE_{min}”より小さくなってしまった場合は、何らかの異常が発生したものとしてエラー(error)情報が照会要求元のユーザ管理サーバ4へ伝送される。

【0104】一方、残りメモリ量に多少余裕のある検索サーバ数“e”が最小分割数“nPE_{min}”以上($e \geq nPE_{min}$)であった場合(ステップS35でYESと判定された場合)は、検索サーバ数“e”のうち、残りメモリ量の大きい検索サーバから最小分割数“nPE_{min}”分だけ検索サーバを抽出して、検索サーバ割当機能58により、それぞれに、検索対象テキストベース8の分割部分81が割り当てられるとともに、テーブル更新機能55により、検索サーバ情報テーブル53における各種の情報が以下のように更新される(ステップS36)。

【0105】・「検索ユーザ数」に“1”加算

・「総検索テキストベース量」に最低メモリ量“c”を加算

・「使用可能メモリ量」から最低メモリ量“c”を減算

・「検索テキストベース名」を登録

・「分割数」に“nPE_{min}”を登録

・「データ量」に最低メモリ量“c”を登録

その後、検索管理サーバ5は、上述のごとく検索サーバ割当機能58により検索対象テキストベース8の分割部分81に対する検索を割り当てられた検索サーバに対して、検索対象テキストベース8の分割部分81とその関連データとをロードするためのローディング指示が、ローディング指示機能59によって、TB管理テーブル51で管理されているテキストファイル名及び関連ファイル名とともに割当先の各検索サーバ6へ送られる(ステップS37)。

【0106】そして、検索管理サーバ5は、AD割当機能56により、AD管理テーブル52に空き項目を1つ割り当て、「参照ユーザ数」=1、対応するテキストベース名、検索サーバID群(検索対象テキストベース8を割り当てられた検索サーバ(群)6に関するローディング情報)を設定したのち(ステップS38)、ローディング指示機能(ローディング情報転送機能)59により、この検索サーバID群とADとを検索サーバ群情報として照会要求元のユーザ管理サーバ4に伝送(転送)

する(ステップS39)。

【0107】つまり、この検索管理サーバ5は、データ分割数決定機能57により決定された分割数“nPE”分の検索サーバを検索サーバ割当機能58により確保できない場合、データ分割数決定機能57により決定された検索対象テキストベース8の分割数“nPE”を見直し、見直した分割数“nPE_{min}”に応じて検索サーバ割当機能58、ローディング指示機能(ローディング情報転送機能)59による処理を行なうようになっているのである。

【0108】従って、検索対象のテキストベース8のサイズが大きい場合でも、常に、各検索サーバ6にその使用可能メモリ量(残りメモリ量)を考慮したサイズのテキストベース8の分割部分81を割り当てることができるので、テキストベース8のサイズが大きすぎるために大幅に検索速度が低下したり、検索不能になったりすることがない。

【0109】なお、上述の検索管理サーバ5では、テキストベース8の分割部分81とその関連データとが、検索対象のテキストベース8の分割数“nPE”と各検索サーバ6の使用可能メモリ量とに応じて、使用可能メモリ量の大きい検索サーバ6から順に割り当てられているが、上述の分割数“nPE”と各検索サーバの負荷状況とに応じて、負荷の軽い検索サーバから順に割り当てるようにしてもよい。

【0110】(c1) 検索管理サーバ5の動作の変形例の説明

次に、以下では、上述の検索管理サーバ5の動作の変形例について、図18、図19に示す各フローチャート

(ステップS41~S53)を用いて説明する。まず、図18に示すように、この場合も、例えば、ユーザ管理サーバ4から検索対象テキストベース8の照会要求があると(ステップS41でYESと判定されると)、検索管理サーバ5では、AD割当機能56によって、AD管理テーブル52に照会要求のあった検索対象テキストベース8が既に存在している(登録されている)かどうか

が判定される(ステップS42)。
【0111】この結果、AD管理テーブル52に検索対象テキストベース8が登録されておらず、検索サーバ6に検索対象テキストベース8が未だロードされていない場合(ステップS42でNOと判定された場合)は、図16、図17により前述したステップS25~S39と同様の処理が行なわれるが、既に検索対象テキストベース8がロードされている場合(ステップS42でYESと判定された場合)は、図19に示すように、データ分割数決定機能57によってTB管理テーブル51からテキストベース8に関する情報が特定され、その中のテキストベース8の実体を格納しているファイルのサイズ情報より、情報検索に必要な総メモリ量“a”が算出される(ステップS43)。

【0112】さらに、データ分割数決定機能57は、検索対象テキストベース8に対応するTB管理テーブル51の使用検索アルゴリズム情報に基づいて、1PE当たりの推奨メモリ量“b”を算出し(ステップS44)、検索対象テキストベース8の分割数“nPE”を、 $nPE = (a + b - 1) / b$ なる演算により算出する(ステップS46)。

【0113】一方、このとき検索サーバ割当機能58では、検索サーバ情報テーブル53から使用可能メモリ量が所定値(ここでは、推奨メモリ量“b”の3倍)以上残っている検索サーバ数“d”を抽出しており(ステップS46)、この検索サーバ数“d”と検索対象テキストベース8の分割数“nPE”とを比較することによって検索サーバ数“d”が分割数“nPE”以上($d \geq nPE$)であるかが判定される(ステップS47)。

【0114】この結果、 $d \geq nPE$ であった場合、つまり、複数の検索サーバの中から推奨メモリ量“b”の3倍よりも使用可能メモリ量の大きい検索サーバを分割数“nPE”だけ確保できる場合は、検索サーバ割当機能58が、検索サーバ数“d”の中から“ $3 \times b$ ”よりも使用可能メモリ量の大きい検索サーバを分割数“nPE”分だけ選択して、検索サーバ情報テーブル53の各種の情報を以下のように更新することにより、検索対象テキストベース8とその関連データとが選択された各検索サーバ6に割り当てられるとともに、検索サーバ情報テーブル53における各種の情報が、テーブル更新機能55により以下のように更新される(ステップS47のYESルートからステップS48)。

【0115】・「検索ユーザ数」に“1”加算
・「総検索テキストベース量」に推奨メモリ量“b”を加算

・「使用可能メモリ量」から推奨メモリ量“b”を減算
・「検索テキストベース名」を登録
・「分割数」に“nPE”を登録
・「データ量」に推奨メモリ量“b”を登録

さらに、検索管理サーバ5は、上述の検索サーバ割当機能58による割当結果に基づいて、検索対象テキストベース8の分割部分81とその関連データとを割当先の各検索サーバへロードするためのTBローディング指示を、ローディング指示機能59により、TB管理テーブル51で管理されているテキストファイル名、関連ファイル名及び使用検索アルゴリズム情報とともに割当先の各検索サーバへ送る(ステップS49)。

【0116】その後、検索管理サーバ5は、AD割当機能56により、AD管理テーブル52に空き項目を1つ割り当て、そこに、「参照ユーザ数」=1、対応するテキストベース名、ローディング先の検索サーバID群などの各パラメータを設定したのち(ステップS50)、ローディング指示機能(ローディング情報転送機能)59により、この検索サーバID群やADなどのテキスト

ベース8を割り当てられた検索サーバ(群)6に関するローディング情報を、検索対象テキストベース8に対する情報検索を実際に行なう検索サーバ群情報としてユーザ管理サーバ4へ転送する(ステップS51)。

【0117】ところで、上述のステップS47において、 $d < nPE$ であった場合、つまり、複数の検索サーバ6の中から推奨メモリ量“b”の3倍よりも使用可能メモリ量の大きい検索サーバ6を分割数“nPE”だけ確保できない場合は、AD割当機能56によって、指定された(照会要求のあった)検索対象テキストベース8に対応するAD管理テーブル52の「参照ユーザ数」が更新(増加)される(ステップS47のNORootからステップS52)。

【0118】そして、既に登録済みでAD管理テーブル52で管理されているテキストベース8についてのローディング情報(既にテキストベース8をロードしている検索サーバ6のID群)が、そのままテキストベース8に対する情報検索を担当している検索サーバ群情報として、ローディング指示機能(ローディング情報転送機能)59によりユーザ管理サーバ4へ転送される(ステップS53)。

【0119】このように、上述の検索管理サーバ5では、照会要求のあった検索対象テキストベース8が既にロードされている場合でも、各検索サーバ6の中から所定値(推奨メモリ量“b”の3倍)よりも使用可能メモリ量の大きい検索サーバをテキストベース8の分割数“nPE”だけ確保できれば、既にロードされているテキストベース8を、所定値よりも使用可能メモリ量の大きい他の検索サーバにも分散して割り当てるので、さらに検索対象のテキストベース8に対する情報検索処理を高速に行なうことができる。

【0120】なお、この場合、既にロードされている検索対象テキストベース8は、上述のように所定値よりも使用可能メモリ量の大きい検索サーバに割り当てているが、所定値よりも負荷の軽い検索サーバに割り当ててもよい。

(d) 検索サーバ6の動作説明

次に、以下では、検索サーバ6の動作について、図20に示すフローチャート(ステップS55～S62)を用いて説明する。

【0121】まず、各検索サーバ6では、図20に示すように、外部から受け付けた要求が、検索管理サーバ5からのTBローディング指示であるか、ユーザ管理サーバ4からの検索要求あるいはマッチデータ回復要求(検索結果の転送要求)であるかが検索サーバ命令解析部62で判定される(ステップS55～S57)。そして、例えば、受け付けた要求が検索管理サーバ5からのTBローディング指示であった場合(ステップS55でYESと判定された場合)は、ファイルアクセス機能63により、TB/Index管理テーブル64(図10参照)のAD

に対応するリストに、このTBローディング指示に伴って送られてくるテキストファイル名、関連ファイル名、使用可能メモリ量、使用検索エンジン(検索アルゴリズム情報)などが登録される(ステップS58)。

【0122】さらに、上述のごとくTBローディング指示を受けた検索サーバ6は、ファイルアクセス機能63により、受け付けたTBローディング指示に伴うテキストファイル名及び関連ファイル名によってそれぞれ指定されるテキストファイル71'及び関連ファイル72'をオープンするとともに、使用可能メモリ量分のテキスト領域65を確保(アロケーション:allocate)し、このテキスト領域65にテキストファイル71'及び関連ファイル72'をロードする(ステップS59)。

【0123】その後、例えば、ユーザ管理サーバ4から上述のごとくロードした検索対象テキストベース8に対する検索要求があった場合(ステップS55でNOと判定され且つステップS56でYESと判定された場合)、この検索要求を受けた検索サーバ6では、検索機能66が、TB/Index管理テーブル64に登録されている検索エンジン61-1～61-m、つまり、検索管理サーバ5からの検索アルゴリズム情報により指定される検索エンジン61-1～61-mを用いて、テキスト領域65に読み込んだテキストベース8の分割部分81に対して、ユーザから指示された検索コマンドでの情報検索を行なう(ステップS60)。

【0124】そして、この検索機能66による検索結果(マッチデータのテキスト領域65内でのアドレス情報)は、検索マッチ情報格納機能67により、順次、マッチポイントテーブル68に格納されてゆき、情報検索が終了した時点で、検索結果送出機能69により、マッチポイントテーブル68に格納されているマッチポイントデータ(アドレス情報)が検索結果として検索要求元のユーザ管理サーバ4に返される(ステップS61)。

【0125】さらに、検索サーバ6は、その後、例えばユーザ管理サーバ4から上述のごとく検索を行なったテキストベース8に対するマッチデータ回復要求を受けると(ステップS55、S56でNOと判定され且つステップS57でYESと判定されると)、検索サーバは、原データ転送機能70により、ADにて特定されるテキストベース8に関してユーザから送られてくるマッチポイントデータ(アドレス情報)を基に、テキスト領域65に読み込んだテキストベース8から該当する原データを取り出して、要求元のユーザ管理サーバ4に転送する(ステップS62)。

【0126】このように、各検索サーバ6は、検索管理サーバ5によって割り当てられた自己が担当するテキストベース8の分割部分81に対してのみ情報検索を行なうので、テキストベース8に対する情報検索の並列・独立処理が容易に実現され、検索対象テキストベースに対する情報検索を極めて高速に行なうことができる。ま

た、この各検索サーバは、検索管理サーバ5からの検索アルゴリズム情報により指定される検索エンジン61-1~61-mのいずれかを用いて、読み込んだテキストベース8の分割部分81に対する情報検索を行なうので、検索対象のテキストベース8に最適な検索エンジンを検索エンジン61-1~61-mから選択してテキストベース8に対する情報検索を行なうことができる。従って、例えば、情報検索の際に、検索精度を重視するか、検索速度を重視するかなどの検索条件に応じて、検索エンジン61-1~61-mを選択して用いて情報検索処理を行なうことができ、情報検索処理の柔軟性にも大いに寄与する。

【0127】以上のように、本実施形態における情報検索システムによれば、検索管理サーバ5が、検索対象テキストベース8とこのテキストベース8についての関連データとをそれぞれ分割して、各検索サーバ6のうちの全部または一部に割り当て、各検索サーバ6が、それぞれ、割り当てられたテキストベース8の分割部分81に対する情報検索を並列かつ独立に行なうので、テキストベース8のサイズが大きくても、各検索サーバ6による情報検索処理の負荷は非常に軽くなる。

【0128】従って、複数のユーザからの検索要求などにより複数のテキストベース8に対する情報検索を同時に行なう場合でも、特別なハードウェアを必要とせず、極めて高速にその処理を行なうことができる。なお、上述の情報検索システムでは、ユーザ管理サーバ4をそなえて、これらのユーザ管理サーバ4がユーザからの各種の要求を受け付けて、その要求に応じた処理を検索管理サーバ5、各検索サーバ6に要求することによって、情報検索処理が実行されるようになっているが、各ユーザ管理サーバ4は、システム構築上の必須要件ではなく、他のサーバからの要求、あるいはユーザからの直接の要求に応じて情報検索処理を実行するようにしてもよい。

【0129】また、各検索サーバ6は、複数種類の検索エンジン61-1~61-mを選択して用いるようになっているが、1種類の検索エンジンのみをそなえてこれを固定的に用いるようにしてもよいし、このような検索エンジンを用いずに検索機能66が、直接、検索対象テキストベース8に対する情報検索を行なうようにしてもよい。

【0130】

【発明の効果】以上詳述したように、本発明の情報検索システムによれば、検索管理サーバが、検索対象テキストベースとこのテキストベースについての関連情報とをそれぞれ分割して、複数の検索サーバのうちの全部または一部に割り当て、複数の検索サーバが、それぞれ、割り当てられたテキストベースの分割部分に対する情報検索を並列かつ独立に行なうので、検索対象テキストベースのサイズが大きくても、各検索サーバによる情報検索

処理の負荷は非常に軽くなる。従って、複数のユーザからの検索要求などにより複数のテキストベースに対する情報検索を同時に行なう場合でも、特別なハードウェアを必要とせず、極めて高速にその処理を行なうことができる（請求項1）。

【0131】具体的に、このとき検索管理サーバは、テキストベースの分割数と各検索サーバの負荷状況とに応じて、テキストベースとその関連情報とを、負荷の軽い（使用可能メモリ量の大きい）検索サーバから順に割り当てるので、特定の検索サーバの負荷が大きくなり過ぎるようなことはなく、各検索サーバは、最適な負荷状況で高速に情報検索処理を行なうことができる（請求項2, 4, 8, 13）。

【0132】一方、各検索サーバは、検索管理サーバからのローディング指示に伴うテキストファイル名および関連情報ファイル名によってそれぞれ指定されるテキストベースの分割部分およびその関連情報を読み込んで情報検索を行なう、つまり、検索管理サーバに割り当てられた自己が担当するテキストベースの分割部分に対してのみ情報検索を行なうので、上述のような情報検索の並列・独立処理を容易に実現して、検索対象テキストベースに対する情報検索を極めて高速に行なうことができる（請求項3, 12）。

【0133】なお、各検索サーバは、検索管理サーバからの検索アルゴリズム情報により指定される検索アルゴリズムを用いて、読み込んだテキストベースの分割部分に対する情報検索を行なうこともできるので、検索対象のテキストベースに最適な検索アルゴリズムを選択してテキストベースに対する情報検索を行なうことができる。従って、例えば、情報検索の際に、検索精度を重視するか、検索速度を重視するかなどの検索条件に応じた検索処理を行なうことができ、情報検索処理の柔軟性にも大いに寄与する（請求項5, 13, 17）。

【0134】また、ユーザプロセスからの検索要求を受け付けるユーザ管理サーバをそなえた場合、このユーザ管理サーバが、ユーザプロセスからの検索要求に応じて、検索管理サーバに、検索対象テキストベースに対する情報検索を行なっている検索サーバ群情報の照会することによって、常に、実際の検索要求を検索対象テキストベースを保持している検索サーバに直接的に行なうことができるので、ユーザプロセスからの検索要求に対して遅延なく情報検索処理を実行することができる（請求項6, 7）。

【0135】さらに、検索管理サーバは、検索対象テキストベースのサイズ情報と各検索サーバの負荷状況とに基づいて決定された分割数分の検索サーバを確保できない場合には、そのテキストベースの分割数を見直して、再度、各検索サーバへのテキストベースの割当処理などを行なうので、検索対象テキストベースのサイズが大きい場合でも、常に、各検索サーバにその使用可能メモリ

量（又は、負荷状況）を考慮したサイズのテキストベース（分割部分）を割り当てることができる。従って、検索対象テキストベースのサイズが大きすぎるために大幅に検索速度が低下したり、検索不能になったりすることがない（請求項9, 14）。

【0136】また、検索管理サーバは、照会要求のあった検索対象テキストベースが既にロードされている場合、検索対象テキストベースについてのローディング情報（検索対象テキストベースを保持している検索サーバ群情報）を、照会要求元のユーザ管理サーバへ転送するので、新たに、検索対象テキストベースを保持している検索サーバを特定しなくてもよく、照会結果に基づいた実際の検索要求を各検索サーバへ出すまでの処理時間を大幅に短縮して高速化することができる（請求項10, 15）。

【0137】さらに、検索管理サーバは、照会要求のあった検索対象テキストベースが既にロードされている場合、複数の検索サーバの中から所定値よりも使用可能メモリ量の大きい（又は、負荷の軽い）検索サーバを検索対象テキストベースの分割数だけ確保できれば、ロードされている検索対象テキストベースを、確保した所定値よりも使用可能メモリ量の大きい（又は、負荷の軽い）検索サーバに分散して割り当てることもできるので、さらに各検索サーバによる情報検索処理を高速に行なうことができる（請求項11, 16）。

【図面の簡単な説明】

【図1】本発明の原理ブロック図である。

【図2】本発明の一実施形態としての情報検索システムの構成を示すブロック図である。

【図3】本実施形態におけるユーザ管理サーバの詳細構成を示すブロック図である。

【図4】本実施形態におけるユーザ固有情報テーブルの一例を示す図である。

【図5】本実施形態における検索管理サーバの詳細構成を示すブロック図である。

【図6】本実施形態におけるテキストベース（TB）管理テーブルの一例を示す図である。

【図7】本実施形態におけるアクセス識別子（AD）管理テーブルの一例を示す図である。

【図8】本実施形態における検索サーバ情報テーブルの一例を示す図である。

【図9】本実施形態における検索サーバの詳細構成を示すブロック図である。

【図10】本実施形態におけるテキストベース／インデックス（TB/Index）管理テーブルの一例を示す図である。

【図11】本実施形態におけるユーザ管理サーバの動作を説明するためのフローチャートである。

【図12】本実施形態におけるユーザ管理サーバの動作を説明するためのフローチャートである。

【図13】本実施形態におけるユーザ管理サーバの動作

を説明するためのフローチャートである。

【図14】本実施形態におけるユーザ管理サーバの動作を説明するためのフローチャートである。

【図15】本実施形態における検索管理サーバの動作を説明するためのフローチャートである。

【図16】本実施形態における検索管理サーバの動作を説明するためのフローチャートである。

【図17】本実施形態における検索管理サーバの動作を説明するためのフローチャートである。

10 【図18】本実施形態における検索管理サーバの動作の変形例を説明するためのフローチャートである。

【図19】本実施形態における検索管理サーバの動作の変形例を説明するためのフローチャートである。

【図20】本実施形態における検索サーバの動作を説明するためのフローチャートである。

【符号の説明】

1, 6 検索サーバ

2, 5 検索管理サーバ

3, 8 テキストベース

20 4 ユーザ管理サーバ

7 ファイルシステム（データベース）

31, 81 分割部分

41 ユーザ固有情報テーブル

41A ユーザ名（User name）

41B ユーザアドレス（User address）

41C ADポインタ（*AD pointer）

41D マッチポイントテーブルポインタ（*Match point tp）

42 外部命令パケット送受信機能（検索要求受付機能、検索結果転送機能）

43 ユーザ管理サーバ命令解析部

44 ユーザ固有情報登録・更新・照会機能

45 検索サーバ照会機能

46 検索要求機能

47 検索マッチデータ回復機能

51 テキストベース（TB）管理テーブル（テキストベース管理機能）

52 アクセス識別子（AD）管理テーブル（ローディング情報管理機能）

53 検索サーバ情報テーブル（検索サーバ管理機能）

54 検索管理サーバ命令解析部

55 テーブル更新機能

56 AD割当機能

57 データ分割数決定機能

58 検索サーバ割当機能

59 ローディング指示機能

61-1～61-m 検索エンジン

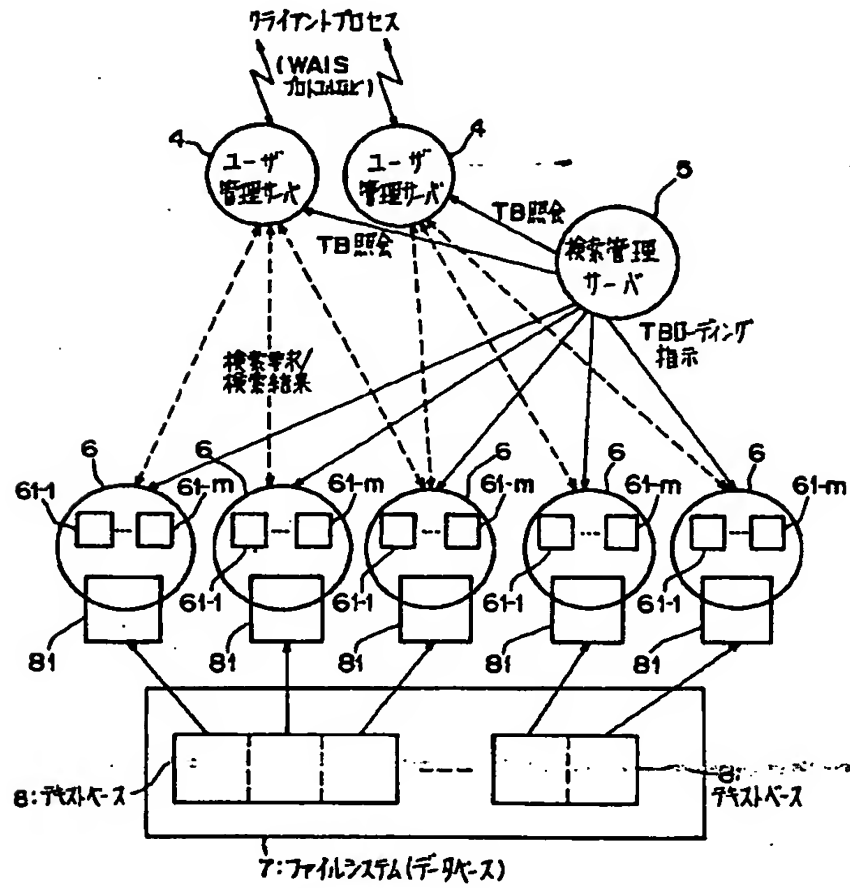
62 検索サーバ命令解析部

63 ファイルアクセス機能

64 テキストベース／インデックス（TB/Index）管理テ

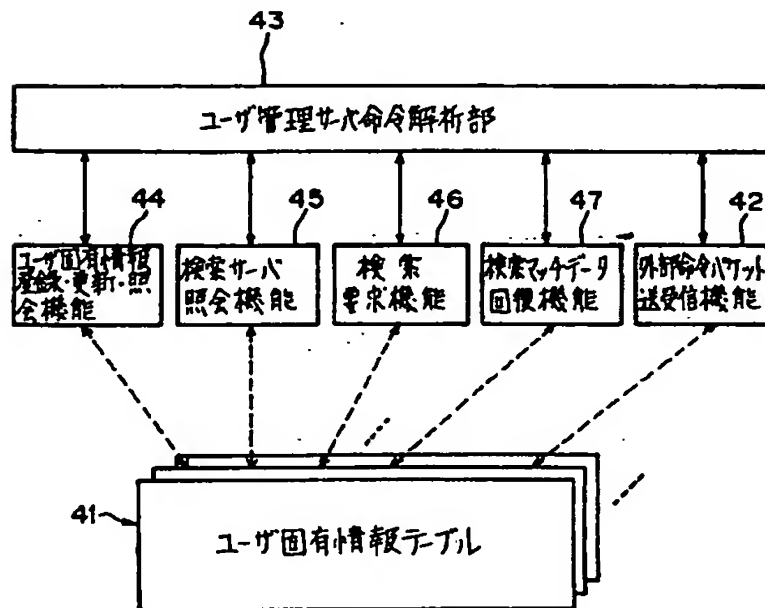
【図2】

本発明の一実施形態としての情報検索システムの構成を示すブロック図



【図3】

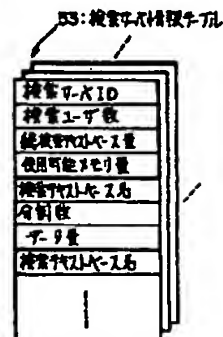
本実施形態におけるユーザ管理サーバの詳細構成を示すブロック図



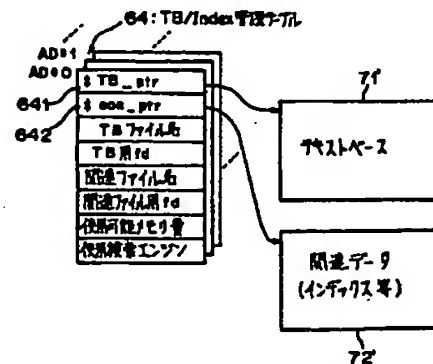
【図8】

【図10】

本実施形態における検索サーバ情報テーブルの一例を示す図

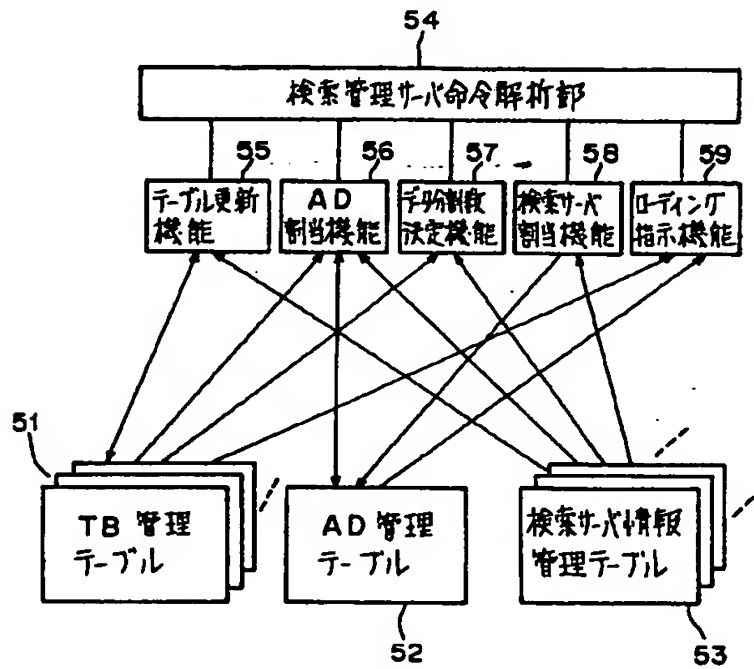


本実施形態におけるファイルシステム/インデックス(TB/Index)管理テーブルの一例を示す図



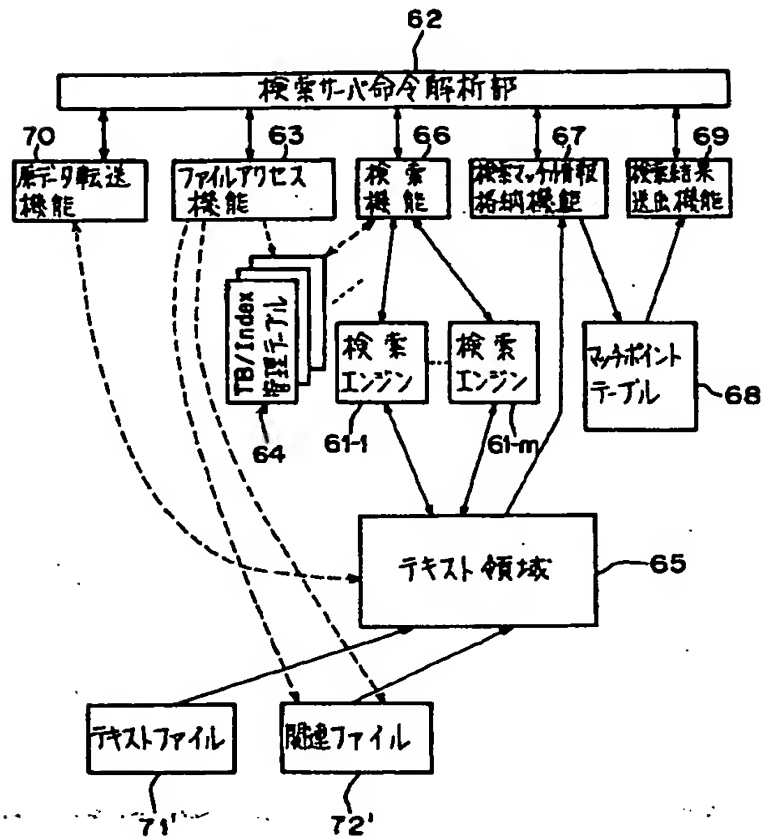
【図5】

本実施形態における検索管理サーバの詳細構成を示すブロック図



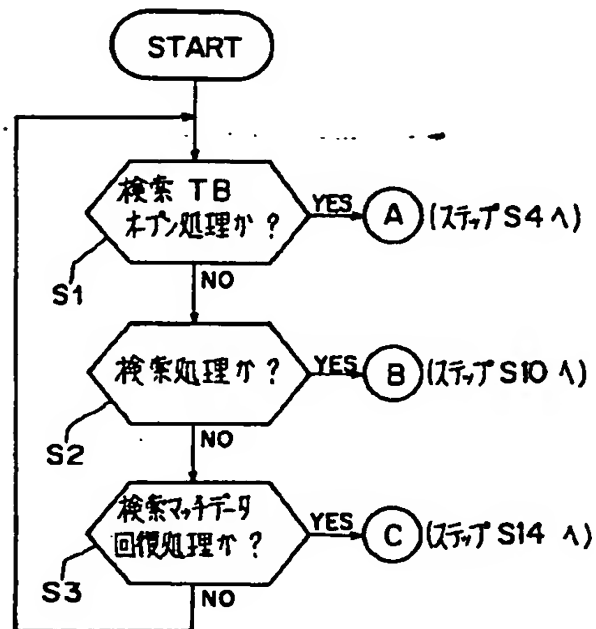
【図9】

本実施形態における検索サーバの詳細構成を示すブロック図



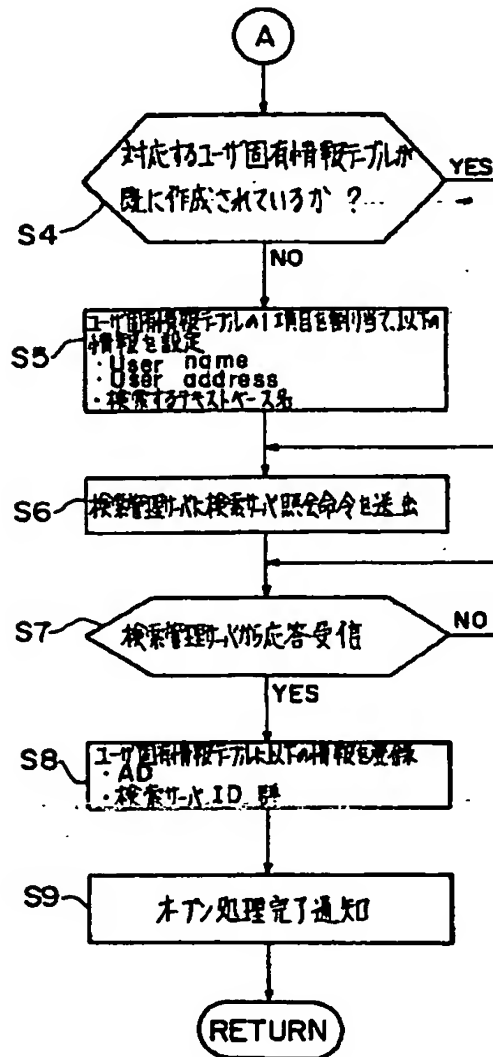
【図11】

本実施形態におけるユーザ管理サーバの動作を説明するためのフローチャート



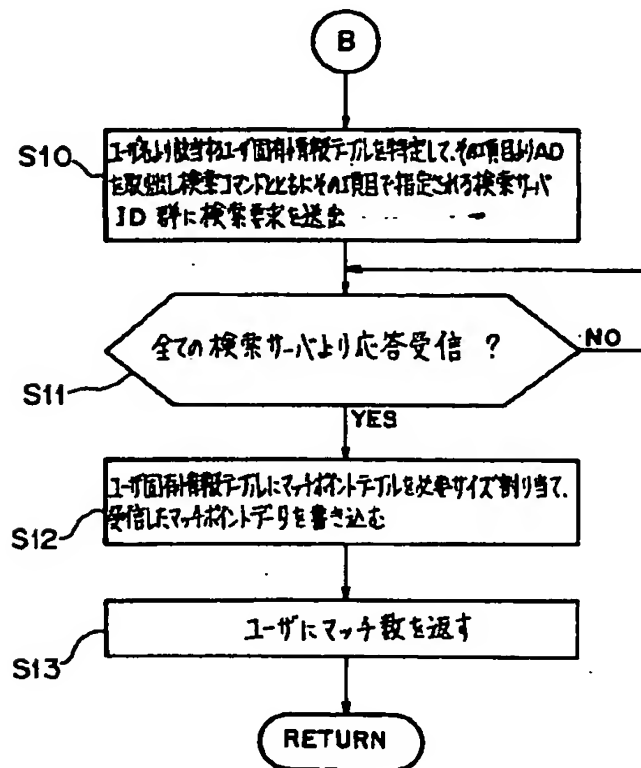
【図12】

本実施形態におけるユーザ管理サーバの動作を説明するためのフローチャート



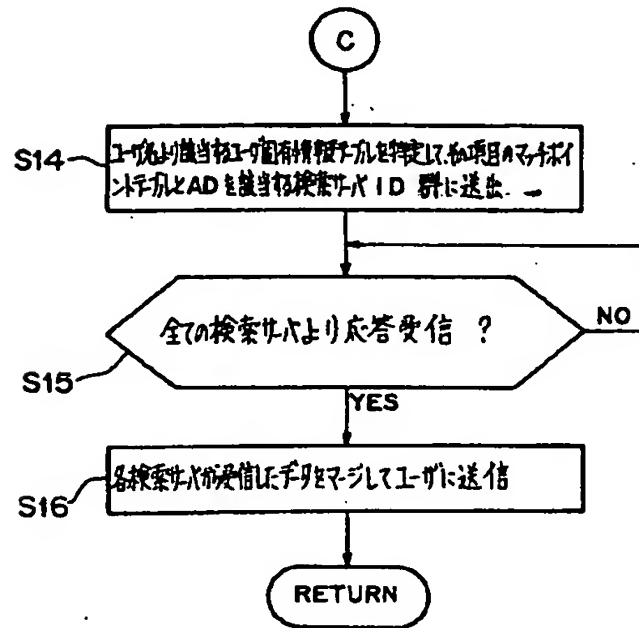
【図13】

本実施形態におけるユーザ管理サーバの動作を説明するためのフローチャート



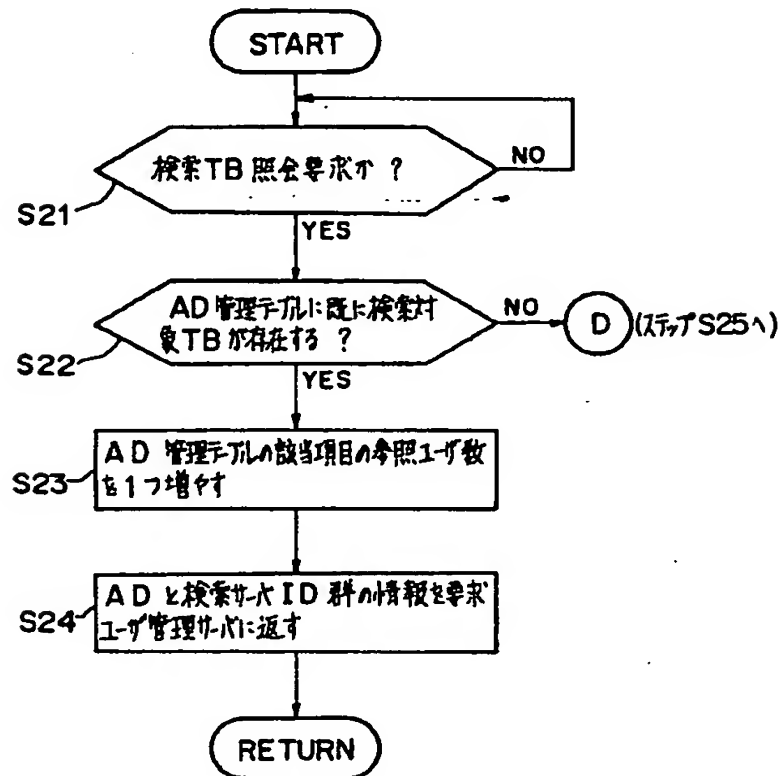
【図14】

本実施形態におけるユーザ管理サーバの動作を説明するためのフローチャート



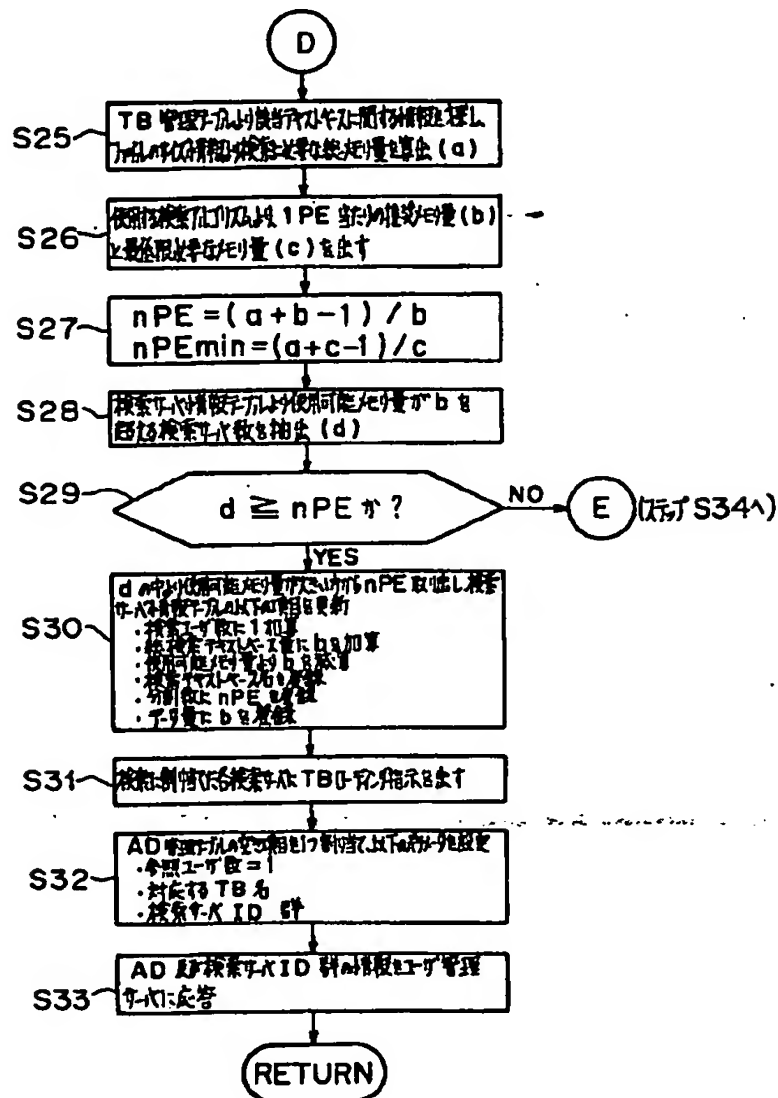
【図15】

本実施形態における検索管理サーバの動作を説明するためのフローチャート



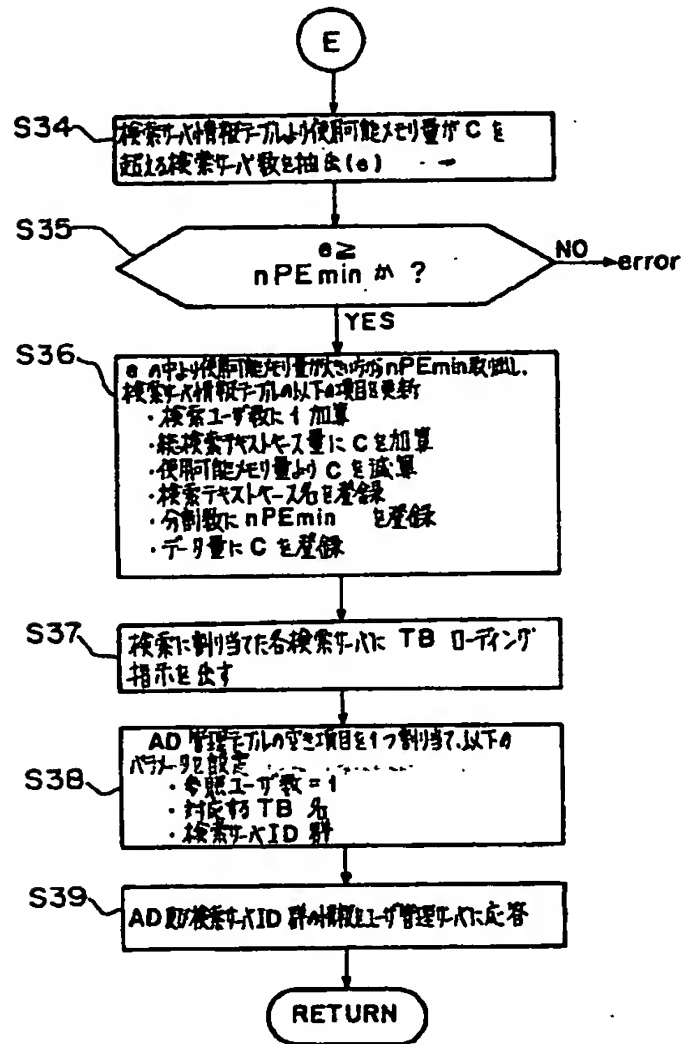
【図16】

本実施形態における検索管理サービスの動作を説明するためのフローチャート



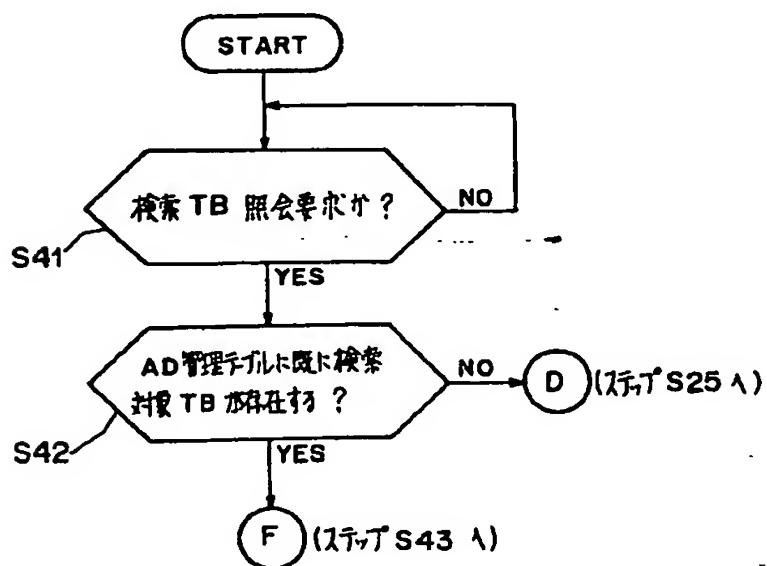
【図17】

本実施形態における検索管理サーバ動作の説明のためのフローチャート



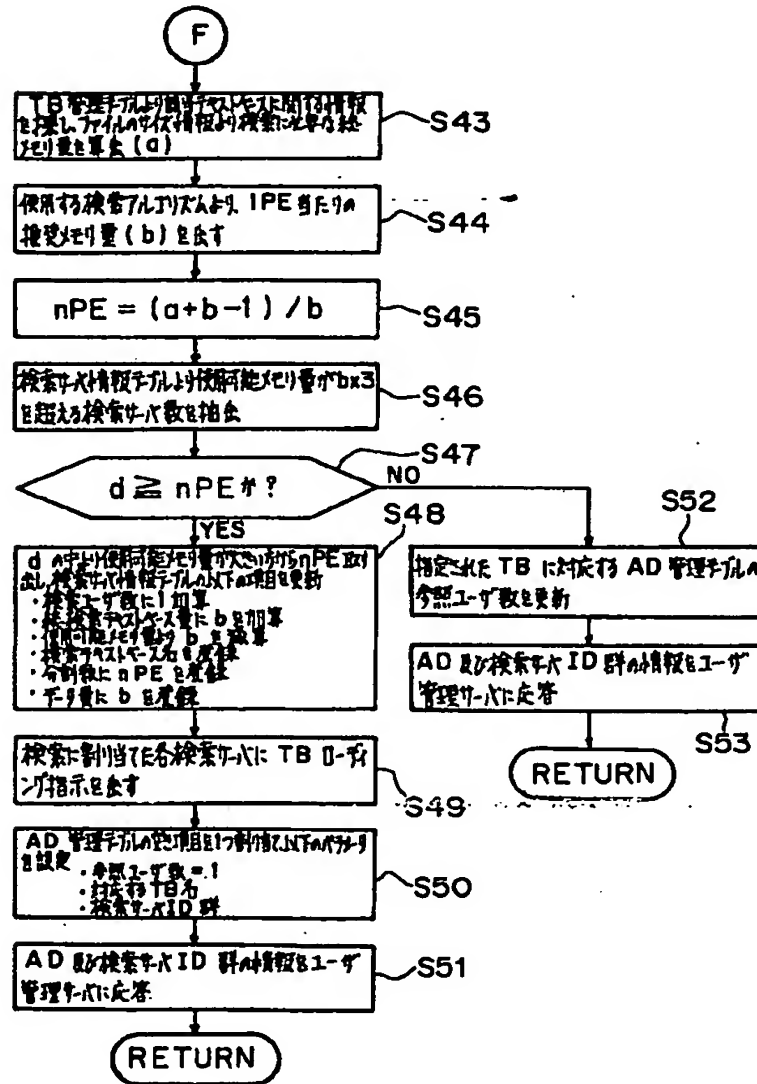
【図18】

本実施形態における検索管理サーバの動作の変形例を説明するためのフローチャート



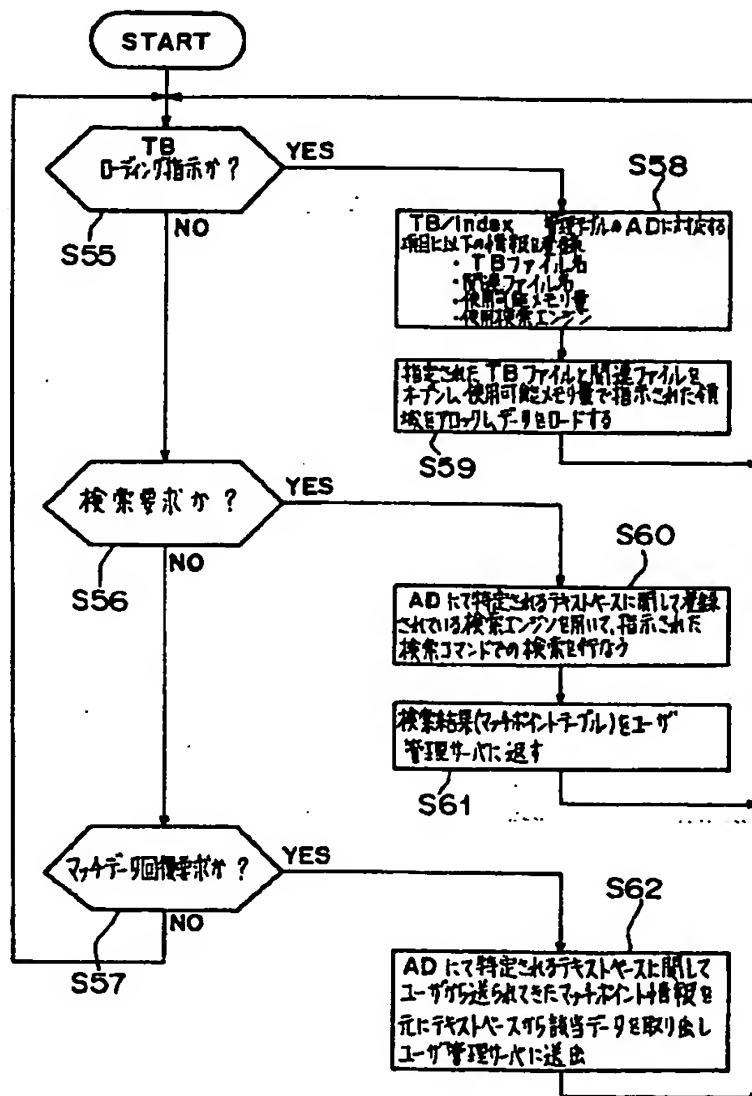
[図19]

本実施形態における検索管理サーバ動作の一例を説明するためのフローチャート



【図20】

本実施形態における検索サーバの動作を説明するためのフローチャート



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.